
May

1

2

•

ein PfsPN

2

P 12/1 pe

2

P 12/1 pe

reted pro

tion-indu

2

oxide unia

peptide unit

peptide units

peptide units

ptide unique

oxide unc

This Page Blank (uspto)

0

Too Late

5

```

AAW32302 Length: 534 May 30, 2002 09:25 Type: P Check: 605
1 MAREOLOVLN ALDVAKTQMT HFTAILIAGM GFETDAVDLF CISLVTKLIG
51 RIYYHVEGAQ KPCTLPNNVA AAVNGVAFCG TLAGOLFPGW LGDKLGRRKV
101 YGMLTMVWVL CSTASGSRFG HEPRKVMATL CFFEFNLGFG IGGDYLSLT
151 IMSEYANKRT KAFVSAVPA MGGFIMAGG IFALIIISAF EAKFPSPAYA
201 DDALGSTRPQ ADLVWRILIM AGAIPAMTY YSRKMPETA RYVALAXDA
251 KQASDSMKV LOYEIEPEQ KLEETSEKS KAFLESEKF MSRHGLHLIG
301 TTSTWFLDII AFYSQNLFOR DIFSAIGWIP PAQSMNAIOE VFKIRAPQTL
351 IALCSTVPGY WTVVAFIDIV GFPAIOMGF PEMVVPFAL AIPYNWTHK
401 ENRIGEVIMY SLTEFFANFG PNATFEVPA EIPPARERST CHGISASGK
451 LGAMVGARGF LYLAQNPDKO KIDAGYPPGI GVRNSLVLG VVNFILFET
501 FLVPESKSGS LEMSGENED NENSNDKRT VPIV

!!!AA-SEQUENCE 1.0
ID AAW32303 standard; Protein: 521 AA.
AC AAW32303;
XX
XX 27-APR-1998 (first entry)
XX
XX Arabidopsis thaliana inorganic phosphate transporter 3.
XX
XX Columbia strain; inorganic phosphate transporter 3; IPN3;
XX accelerated phosphate uptake; tobacco plant.
XX
XX Arabidopsis thaliana.
XX
XX MW9735984-A1.
XX

```

This Page Blank (uspto)

FINDPATTERNS on geneseq: * allowing 0 mismatches		May	
1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	AAW82385
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	ck: 309 len: 355 1 Aaw82385 Flea saliva protein Pfspn6-357.
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P)XFX(R)(F)WXX(O)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	PGFSRFWNEO
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	QCPAY
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	AAW82322
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	ck: 4400 len: 19 1 Aaw82322 p53 homologue TIP 12/1 peptide.
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P)XFX(D)(Y)WXX(L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	PRFMDYWEGL
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	NENG
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	AAW82320
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	ck: 4400 len: 19 1 Aaw82320 p53 homologue TIP 12/1 peptide.
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P)XFX(D)(Y)WXX(L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	PRFMDYWEGL
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	NENG
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	AAW58852
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	ck: 5772 len: 130 1 Aaw58852 Human O289_1 secreted protein.
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(C)XFX(S)(F)WXX(L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	CHRSSFWIGL
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	PA
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	AAW56025
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	ck: 9188 len: 316 1 Aaw56025 Phosphate starvation-induced pr
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(L)XFX(R)(F)WXX(F)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	LCFFRFWIGL
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	GIGGD
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	AAW37220
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	ck: 5978 len: 12 1 Aaw37220 MDW2 binding peptide unique pha
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P)XFX(D)(Y)WXX(L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	PRFMDYWEGL
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	N
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	AAW37224
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	ck: 9428 len: 15 1 Aaw37224 MDW2 binding peptide unique pha
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(L)XFX(D)(Y)WXX(L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	LVEFADYWEGL
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	Y
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	AAW37221
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	ck: 6151 len: 12 1 Aaw37221 MDW2 binding peptide unique pha
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(Q)XFX(D)(Y)WXX(O)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	QNFIDYWTQO
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	F
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	AAW37222
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	ck: 5993 len: 12 1 Aaw37222 MDW2 binding peptide unique pha
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P)XFX(H)(Y)WXX(F)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	PAFTYHWFATF
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	TG
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	AAW37223
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	ck: 9093 len: 15 1 Aaw37223 MDW2 binding peptide unique pha
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P)XFX(D)(H)WXX(L)
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	PTFRDHWFL
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	V
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	AAW37225
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	ck: 8833 len: 15 1 Aaw37225 MDW2 binding peptide unique pha

1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(L) PAFSRFMSDL	SAGAH	1	2: Q	PTFSDYWKLL	P
1	AAW37226 ck: 4557 len: 10 i Aaw37226 MDM2 binding peptide phage consens (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PFXSDYMXKL		1	2: Q	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(Y)WXX(L) ETFSDYWKLL	P
1	AAW37195 ck: 5993 len: 12 i Aaw37195 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(H)(Y)WXX(F) PAFTHWATF		1	2: V	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(D)(Y)WXX(Q) QNFIDYWTQ	F
1	AAW37196 ck: 5978 len: 12 i Aaw37196 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PRFMDYWEGL	N	1	5: IDRA	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(H)WXX(L) PFRDHWFL	V
1	AAW37197 ck: 8253 len: 14 i Aaw37197 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PTFSDYWKLL	P	1	5: PRPA	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(Y)WXX(L) LVFADYWEFL	Y
1	AAW37203 ck: 634 len: 16 i Aaw37203 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(Y)WXX(L) ETFSDYWKLL	P	1	1:	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(L) PAFSRFMSDL	SAGAH
1	AAW37204 ck: 1571 len: 28 i Aaw37204 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PRFMDYWEGL	NROIK	1	2: M	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PRFMDYWEGL	N
1	AAW37205 ck: 8233 len: 31 i Aaw37205 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PRFMDYWEGL	NXK	1	2: Q	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PTFSDYWKLL	P
1	AAW37182 ck: 5993 len: 12 i Aaw37182 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(H)(Y)WXX(F) PAFTHWATF		1	5: PRPA	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(Y)WXX(L) LVFADYWEFL	Y
1	AAW37183 ck: 5978 len: 12 i Aaw37183 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PRFMDYWEGL	N	1	130: AVMAT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWGLF	GIGGD
1	AAW37184 ck: 6146 len: 12 i Aaw37184 Human oncogenic protein MDM2 bindi (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PRFMDYWEGL		1	AAW46749 ck: 4793 len: 538 i Aaw46749 Tomato phosphate transporter 1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWGLF	GIGGD	
1				129: GVMET		

1	AAW46750	ck: 3154	len: 528	1	Aaw46750	Tomato phosphate transporter 1 prd
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
135:	GVWTT				LCFREWLTGF	GIGGD
1	AAW37145	ck: 5772	len: 130	1	Aaw37145	Human O289_1 protein. 6/1998
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(C)XFX(S)(F)WXX(L)				
119:	VISHL				CHRSSFMTGL	PA
1	AAV30723	ck: 7459	len: 182	1	Aay30723	Amino acid sequence of a human sec
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(H)WXX(F)				
124:	LOPTR				LNKDHMTIF	SVASA
1	AAV31622	ck: 7787	len: 284	1	Aay31622	Human oxidised LDL receptor (HOLR)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(E)XFX(S)(Y)WXX(L)				
195:	SOSYS				EFERYSYMTGL	LRPDS
1	AAV38483	ck: 9545	len: 86	1	Aay38483	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
19:	MKKG				LYSDFWNKL	DVGAI
1	AAV38484	ck: 8317	len: 27	1	Aay38484	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
11:	MKKG				LYSDFWNKL	DVGAI
1	AAV36885	ck: 7481	len: 945	1	Aay36885	Protein involved in intermediate m
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(E)XFX(S)(F)WXX(F)				
674:	RSSKM				ELFLSFWTAF	SLFLP
1	AAV38424	ck: 463	len: 509	1	Aay38424	Human secreted protein. 9/1999
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
103:	MKKA				LYSDFWNKL	DVGAI
1	AAV34694	ck: 6409	len: 945	1	Aay34694	C. pneumoniae protein involved in
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(E)XFX(S)(Y)WXX(F)				
675:	RSSKM				ELFLSYWSGF	IPHLN
1	AAV05317	ck: 2849	len: 280	1	Aay05317	Human secreted protein bn97_1. 6/1
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(E)XFX(S)(Y)WXX(L)				
191:	SOSYS				EFERYSYMTGL	LRPDS
1	AAW67815	ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				
		ck: 5177	len: 57	1	Aaw67815	Human secreted protein encoded by
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(L)				
		(E)XFX(S)(Y)WXX(L)				
		EFERYSYMTGL				
		LRPDS				

1	AAB17078	ck: 6146	len: 12	1	Aab17078	Mdm/hdm antagonist peptide sequenc	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PTFSDYWKLL	P
1	2:	Q						
1	AAB17079	ck: 5978	len: 12	1	Aab17079	Mdm/hdm antagonist peptide sequenc	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PREMDYWEGL	N
1	2:	M						
1	AAB17080	ck: 6151	len: 12	1	Aab17080	Mdm/hdm antagonist peptide sequenc	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(D)(Y)WXX(Q) QNFIDYWTQQ	F
1	2:	V						
1	AAB17081	ck: 5993	len: 12	1	Aab17081	Mdm/hdm antagonist peptide sequenc	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(H)(Y)WXX(F) PAFTHWATF	
1	3:	TG						
1	AAB17082	ck: 9093	len: 15	1	Aab17082	Mdm/hdm antagonist peptide sequenc	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(H)WXX(L) PTFRDHWEAL	V
1	5:	IDRA						
1	AAB17083	ck: 9428	len: 15	1	Aab17083	Mdm/hdm antagonist peptide sequenc	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(Y)WXX(L) LVFADYWEIL	Y
1	5:	PRPA						
1	AAB17084	ck: 8833	len: 15	1	Aab17084	Mdm/hdm antagonist peptide sequenc	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(L) PAFSRFWSDL	SAGAH
1	1:							
1	AAB17085	ck: 8896	len: 15	1	Aab17085	Mdm/hdm antagonist peptide sequenc	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(L) PAFSRFWSKL	SAGAH
1	1:							
1	AAB17086	ck: 4557	len: 10	1	Aab17086	Mdm/hdm antagonist peptide sequenc	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(L) PAFSRFWSKL	SAGAH
1	1:							
1	AAB17089	ck: 6124	len: 12	1	Aab17089	Mdm/hdm antagonist peptide sequenc	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(Y)WXX(L) EFPSDYWKLL	P
1	2:	Q						
1	AAB17090	ck: 6146	len: 12	1	Aab17090	Mdm/hdm antagonist peptide sequenc		
1	2:	Q						
1	191:	PLASD						
1	280:	PLASD						
1	AAG29844	ck: 4529	len: 461	1	Aag29844	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(L) LSFGRYWYTL	RKQTF
1	5419	ck: 5419	len: 504	1	Aag21029	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	530	ck: 530	len: 478	1	Aag17204	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	526:	SAWAM						
1	561:	SAWAM						
1	AAG17203	ck: 1255	len: 549	1	Aag17203	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	561:	SAWAM						
1	AAG17202	ck: 7959	len: 584	1	Aag17202	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(F)WXX(L) LYPSDSWVKLL	DVGAI
1	862:	MKAA						
1	AAY92944	ck: 5055	len: 1,503	1	Aay92944	Human TRPC7 protein. 11/2000	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(F)WXX(L) LYPSDSWVKLL	P
1	2:	Q						
1	191:	PLASD						
1	280:	PLASD						
1	AAG29844	ck: 4529	len: 461	1	Aag29844	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(L) LSFGRYWYTL	RKQTF
1	5419	ck: 5419	len: 504	1	Aag21029	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	530	ck: 530	len: 478	1	Aag17204	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	526:	SAWAM						
1	561:	SAWAM						
1	AAG17203	ck: 1255	len: 549	1	Aag17203	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	561:	SAWAM						
1	AAG17202	ck: 7959	len: 584	1	Aag17202	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(F)WXX(L) LYPSDSWVKLL	DVGAI
1	862:	MKAA						
1	AAY92944	ck: 5055	len: 1,503	1	Aay92944	Human TRPC7 protein. 11/2000	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PTFSDYWKLL	P
1	2:	Q						
1	191:	PLASD						
1	280:	PLASD						
1	AAG29844	ck: 4529	len: 461	1	Aag29844	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(L) LSFGRYWYTL	RKQTF
1	5419	ck: 5419	len: 504	1	Aag21029	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	530	ck: 530	len: 478	1	Aag17204	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	526:	SAWAM						
1	561:	SAWAM						
1	AAG17203	ck: 1255	len: 549	1	Aag17203	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	561:	SAWAM						
1	AAG17202	ck: 7959	len: 584	1	Aag17202	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(F)WXX(L) LYPSDSWVKLL	DVGAI
1	862:	MKAA						
1	AAY92944	ck: 5055	len: 1,503	1	Aay92944	Human TRPC7 protein. 11/2000	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PTFSDYWKLL	P
1	2:	Q						
1	191:	PLASD						
1	280:	PLASD						
1	AAG29844	ck: 4529	len: 461	1	Aag29844	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(L) LSFGRYWYTL	RKQTF
1	5419	ck: 5419	len: 504	1	Aag21029	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	530	ck: 530	len: 478	1	Aag17204	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	526:	SAWAM						
1	561:	SAWAM						
1	AAG17203	ck: 1255	len: 549	1	Aag17203	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	561:	SAWAM						
1	AAG17202	ck: 7959	len: 584	1	Aag17202	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(F)WXX(L) LYPSDSWVKLL	DVGAI
1	862:	MKAA						
1	AAY92944	ck: 5055	len: 1,503	1	Aay92944	Human TRPC7 protein. 11/2000	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PTFSDYWKLL	P
1	2:	Q						
1	191:	PLASD						
1	280:	PLASD						
1	AAG29844	ck: 4529	len: 461	1	Aag29844	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(L) LSFGRYWYTL	RKQTF
1	5419	ck: 5419	len: 504	1	Aag21029	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	530	ck: 530	len: 478	1	Aag17204	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	526:	SAWAM						
1	561:	SAWAM						
1	AAG17203	ck: 1255	len: 549	1	Aag17203	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	561:	SAWAM						
1	AAG17202	ck: 7959	len: 584	1	Aag17202	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(F)WXX(L) LYPSDSWVKLL	DVGAI
1	862:	MKAA						
1	AAY92944	ck: 5055	len: 1,503	1	Aay92944	Human TRPC7 protein. 11/2000	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PTFSDYWKLL	P
1	2:	Q						
1	191:	PLASD						
1	280:	PLASD						
1	AAG29844	ck: 4529	len: 461	1	Aag29844	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(L) LSFGRYWYTL	RKQTF
1	5419	ck: 5419	len: 504	1	Aag21029	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	530	ck: 530	len: 478	1	Aag17204	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	526:	SAWAM						
1	561:	SAWAM						
1	AAG17203	ck: 1255	len: 549	1	Aag17203	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	561:	SAWAM						
1	AAG17202	ck: 7959	len: 584	1	Aag17202	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(F)WXX(L) LYPSDSWVKLL	DVGAI
1	862:	MKAA						
1	AAY92944	ck: 5055	len: 1,503	1	Aay92944	Human TRPC7 protein. 11/2000	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PTFSDYWKLL	P
1	2:	Q						
1	191:	PLASD						
1	280:	PLASD						
1	AAG29844	ck: 4529	len: 461	1	Aag29844	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(L) LSFGRYWYTL	RKQTF
1	5419	ck: 5419	len: 504	1	Aag21029	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	530	ck: 530	len: 478	1	Aag17204	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	526:	SAWAM						
1	561:	SAWAM						
1	AAG17203	ck: 1255	len: 549	1	Aag17203	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	561:	SAWAM						
1	AAG17202	ck: 7959	len: 584	1	Aag17202	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(F)WXX(L) LYPSDSWVKLL	DVGAI
1	862:	MKAA						
1	AAY92944	ck: 5055	len: 1,503	1	Aay92944	Human TRPC7 protein. 11/2000	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PTFSDYWKLL	P
1	2:	Q						
1	191:	PLASD						
1	280:	PLASD						
1	AAG29844	ck: 4529	len: 461	1	Aag29844	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(L) LSFGRYWYTL	RKQTF
1	5419	ck: 5419	len: 504	1	Aag21029	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	530	ck: 530	len: 478	1	Aag17204	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	526:	SAWAM						
1	561:	SAWAM						
1	AAG17203	ck: 1255	len: 549	1	Aag17203	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	561:	SAWAM						
1	AAG17202	ck: 7959	len: 584	1	Aag17202	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(F)WXX(L) LYPSDSWVKLL	DVGAI
1	862:	MKAA						
1	AAY92944	ck: 5055	len: 1,503	1	Aay92944	Human TRPC7 protein. 11/2000	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PTFSDYWKLL	P
1	2:	Q						
1	191:	PLASD						
1	280:	PLASD						
1	AAG29844	ck: 4529	len: 461	1	Aag29844	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(L) LSFGRYWYTL	RKQTF
1	5419	ck: 5419	len: 504	1	Aag21029	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	530	ck: 530	len: 478	1	Aag17204	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	526:	SAWAM						
1	561:	SAWAM						
1	AAG17203	ck: 1255	len: 549	1	Aag17203	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVLL	LSVSF
1	561:	SAWAM						
1	AAG17202	ck: 7959	len: 584	1	Aag17202	Arabidopsis thaliana protein fr	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(F)WXX(L) LYPSDSWVKLL	DVGAI
1	862:	MKAA						
1	AAY92944	ck: 5055	len: 1,503	1	Aay92944	Human TRPC7 protein. 11/2000	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(Y)WXX(L) PTFSDYWKLL	P
1	2:	Q						
1	191:	PLASD						
1	280:	PLASD						

1	AAG36897	ck: 2209	len: 333	1	Aag36897	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	862: MKKAA	LYFSPDMNKL	DVCAI			
1	303: INLSS							(P)XFX(D)(H)WXX(L)								
1	AAG36898	ck: 464	len: 318	1	Aag36898	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	AAY66728	ck: 2849	len: 280	1	Aay66728	Membrane-bound protein PRO1131.
1	288: INLSS							(P)XFX(D)(H)WXX(L)			191: SOSYS					
1	AAG36899	ck: 6630	len: 284	1	Aag36899	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	AAY76218	ck: 8440	len: 161	1	Aay76218	Human secreted protein encoded
1	254: INLSS							(P)XFX(D)(H)WXX(L)			97: GLNVP					
1	AAG37272	ck: 7852	len: 584	1	Aag37272	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	AAY53783	ck: 6982	len: 541	1	Aay53783	A maize phosphate transporter d
1	561: SAMAM							(L)XFX(C)(F)WXX(L)			132: GVMAT					
1	AAG37273	ck: 1025	len: 549	1	Aag37273	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	AAY53784	ck: 8403	len: 542	1	Aay53784	A maize phosphate transporter d
1	526: SAMAM							(L)XFX(C)(F)WXX(L)			132: GVMAT					
1	AAG37274	ck: 969	len: 478	1	Aag37274	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	AAY53785	ck: 6498	len: 547	1	Aay53785	A maize phosphate transporter d
1	455: SAMAM							(L)XFX(C)(F)WXX(L)			130: SVMAT					
1	AAG38413	ck: 3958	len: 261	1	Aag38413	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	AAY53786	ck: 3508	len: 539	1	Aay53786	A maize phosphate transporter d
1	49: FLRKS							(P)XFX(S)(H)WXX(L)			132: GVMAT					
1	AAG38839	ck: 7690	len: 986	1	Aag38839	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	AAY53787	ck: 2070	len: 509	1	Aay53787	A maize phosphate transporter d
1	963: SAMAM							(L)XFX(C)(F)WXX(L)			125: NVAV					
1	AAG38840	ck: 7810	len: 886	1	Aag38840	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	AAY53788	ck: 9398	len: 554	1	Aay53788	A maize phosphate transporter d
1	863: SAMAM							(L)XFX(C)(F)WXX(L)			137: AVIGT					
1	AAG38841	ck: 4753	len: 859	1	Aag38841	Arabidopsis thaliana protein fragm	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	ABB63883	ck: 9003	len: 2,040	1	Abb63883	Drosophila melanogaster polypep
1	836: SAMAM							(L)XFX(C)(F)WXX(L)			749: IFLVL					
1	AAV95439	ck: 5055	len: 1,503	1	Aay95439	Human calcium channel polypeptide	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	ABB69566	ck: 7480	len: 337	1	Abb69566	Drosophila melanogaster polypep
								(L)XFX(C)(F)WXX(L)								

183: YPTAE	EDFDEFMNVL	RIVGF	1	ABB34423	ck: 1841	len: 113	1	Abb34423 Peptide #1929 encoded by human
AAU45398	ck: 183	len: 89	1	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
				(E)xfx(h)(f)wxq(Q)				
70: QRVSP	(P)xfx(R)(H)wx(L)	HEIRA		ESFIHFWTKQ				NPTVL
ABG03782	ck: 3344	len: 146	1	ABBI9835	ck: 1841	len: 113	1	Abbi9835 Protein #1834 encoded by probe
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
				(E)xfx(h)(f)wxq(Q)				
7: KKIDR	(P)xfx(D)(F)wx(F)	LDSLK		70: LCEIL				NPTVL
				ESFIHFWTKQ				
ABG03783	ck: 8253	len: 244	1	AAB86968	ck: 4326	len: 506	1	Aab86968 D. melanogaster peptide recepto
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
				(E)xfx(e)(f)wx(L)				
103: KKIDR	(P)xfx(D)(F)wx(F)	LDSVK		323: YPTAE				RIVGF
				EDFDEFMNVL				
ABG05207	ck: 5356	len: 636	1	AAE11926	ck: 2953	len: 288	1	Aae11926 Human CG27 (or C869) receptor p
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
				(E)xfx(s)(Y)wx(L)				
487: SPFER	QLFVCYNALL	ETEHL		199: SQSYS				LRPDS
ABG05563	ck: 5356	len: 636	1	AAE11932	ck: 2849	len: 280	1	Aae11932 Human CG27 (or C869) receptor p
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
				(E)xfx(s)(Y)wx(L)				
487: SPFER	QLFVCYNALL	ETEHL		191: SQSYS				LRPDS
ABG06670	ck: 936	len: 236	1	AAE11933	ck: 6095	len: 314	1	Aae11933 Human CG27 (or C869) receptor p
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
				(E)xfx(s)(Y)wx(L)				
38: GVLWY	LTFSFWAPL	RSPVP		218: SQSYS				LRPDS
ABG06767	ck: 5356	len: 636	1	AAE11934	ck: 89	len: 247	1	Aae11934 Human CG27 (or C869) receptor p
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
				(E)xfx(s)(Y)wx(L)				
487: SPFER	QLFVCYNALL	ETEHL		158: SQSYS				LRPDS
ABG17402	ck: 5356	len: 636	1	AAE11938	ck: 2849	len: 280	1	Aae11938 Human 1lipid metabolism related
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
				(E)xfx(s)(Y)wx(L)				
487: SPFER	QLFVCYNALL	ETEHL		191: SQSYS				LRPDS
ABG29764	ck: 9364	len: 198	1	AAE11939	ck: 2921	len: 307	1	Aae11939 Human 1lipid metabolism related
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
				(E)xfx(s)(Y)wx(L)				
59: GVTRP	EDFVHFWHRF	RTKGD		218: SQSYS				LRPDS
ABB29544	ck: 8372	len: 27	1	AAE11943	ck: 7726	len: 281	1	Aae11943 Human CG27 (or C869) receptor p
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
				(E)xfx(s)(Y)wx(F)				
8: KKIDR	PLFRDFWQOF	LDSLK		185: SQSYS				LRPDS
				EFFSYMTGL				
				AAU20199	ck: 7019	len: 85	1	Aau20199 Human novel endocrine antigen,
				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				

42: GSEET	(L)XFX(S)(H)WXX(Q) LLEFISHWPEQ	SEMA1	
AAU27511	ck: 6166 len: 189 1 Aau27511 Human G-Protein Coupled Receptor (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(S)(Y)WXX(Q) LCFMSYWFPEQ		1
167: SYVRV		FFPGG	
AAU29324	ck: 2849 len: 280 1 Aau29324 Human PRO polypeptide sequence #30 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(Y)WXX(L) EFFYSYWTGL	LRPDS	1
191: SOSYS			
AAU33368	ck: 4809 len: 273 1 Aau33368 Novel human secreted protein #2859 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLFKDFWQOF	LDSYK	1
88: KKIDR			
AAM55526	ck: 8372 len: 27 1 Aam55526 Human brain expressed single exon (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLFKDFWQOF	LDSLK	1
8: KKIDR			
AAM67910	ck: 8372 len: 27 1 Aam67910 Human bone marrow expressed probe (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLFKDFWQOF	LDSLK	1
8: KKIDR			
AAM79021	ck: 5679 len: 377 1 Aam79021 Human protein SEQ ID NO 1683. 11/2 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(F) LFFGHHWOLF	NALTL	1
279: FLLEP			
AAM80005	ck: 7978 len: 406 1 Aam80005 Human protein SEQ ID NO 3651. 11/2 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(F) LFFGHHWOLF	NALTL	1
342: FLLEP			
AAM84808	ck: 1969 len: 36 1 Aam84808 Human Immune/haematopoietic antigen (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (C)XFX(S)(F)WXX(F) CCFSSEFWKLF	RTGHL	1
15: GLIFC			
AAM93544	ck: 2849 len: 280 1 Aam93544 Human polypeptide, SEQ ID NO: 3297 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(Y)WXX(L) EFFYSYWTGL	LRPDS	1
191: SOSYS			
AAO01590	ck: 9232 len: 113 1 Aao01590 Human polypeptide SEQ ID NO 15482. (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(C)(F)WXX(L) PGFPCFWAPL	KKFPP	1
59: FLNPP			
AAB47415	ck: 6481 len: 585 1 Aab47415 Plant phosphate transporter, Le (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFRRWLGF	GIGGD	1
130: GVIGT			
AAM15727	ck: 8372 len: 27 1 Aam15727 Peptide #2161 encoded by probe (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLFKDFWQOF	LDSLK	1
8: KKIDR			
AAM28232	ck: 8372 len: 27 1 Aam28232 Peptide #2269 encoded by probe (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLFKDFWQOF	LDSLK	1
8: KKIDR			
AAM39779	ck: 8327 len: 80 1 Aam39779 Human polypeptide SEQ ID NO 292 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLFKDFWQOF	LDSLK	1
17: GLNVP			
AAM39876	ck: 4914 len: 531 1 Aam39876 Human polypeptide SEQ ID NO 302 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(L) PYELRWETL	DRYMY	1
509: RHLDK			
AAM40139	ck: 2511 len: 258 1 Aam40139 Human polypeptide SEQ ID NO 328 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLFKDFWQOF	LDSLK	1
73: KKIDR			
AAM41565	ck: 4766 len: 104 1 Aam41565 Human polypeptide SEQ ID NO 649 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(F) LLEFYHFWRYF	HCPAD	1
41: GLNVP			
AAM41925	ck: 2688 len: 207 1 Aam41925 Human polypeptide SEQ ID NO 685 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLFKDFWQOF	LDSLK	1
22: KKIDR			
AAU12400	ck: 2849 len: 280 1 Aau12400 Human PRO1131 polypeptide sequ (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(Y)WXX(L) EFFYSYWTGL	LRPDS	1
191: SOSYS			
AAG92957	ck: 4387 len: 592 1 Aag92957 C glutamicum protein fragment S (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(Y)WXX(L) EERIDFWAL	LDRYV	1
559: GHNYH			
AAG92978	ck: 2649 len: 507 1 Aag92978 C glutamicum protein fragment S		

```

1      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
348: EAAEG      (E)XFX(R)(F)WXX(L)      KSRRD
      AAG74295 ck: 942 len: 66      1 Aag74295 Human colon cancer antigen protein
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
40: QCMHE      (L)XFX(H)(H)WXX(Q)      SWXNL
      LFFMHMGCAQ
1      AAG89196 ck: 4447 len: 160      1 Aag89196 Human secreted protein, SEQ ID NO:
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
97: GLNVP      (L)XFX(H)(F)WXX(F)      HCPAD
      LLFFHFMRYP
1      AAU03216 ck: 0 len: 483      1 Auu03216 Fruit fly G protein coupled recept
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
323: YPTAE      (E)XFX(E)(F)WXX(L)      RILPK
      EDPEFWMNVL
1      AAB93312 ck: 4753 len: 732      1 Aab93312 Human protein sequence SEQ ID NO:1
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
547: KKIDR      (P)XFX(D)(F)WXX(F)      LDSLK
      PLFDFWQRF
1      AAB94345 ck: 6460 len: 561      1 Aab94345 Human protein sequence SEQ ID NO:1
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
376: KRIDR      (P)XFX(D)(F)WXX(F)      LDSLK
      PLFDFWQRF
1      AAB60464 ck: 4282 len: 160      1 Aab60464 Human cell cycle and proliferation
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
97: GLNVP      (L)XFX(H)(F)WXX(F)      HCPAD
      LFFHFMRYP
1      AAB60465 ck: 4914 len: 531      1 Aab60465 Human cell cycle and proliferation
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
509: RHIDR      (P)XFX(R)(F)WXX(L)      DRYMV
      PVFLRFWETL
1      AAB65251 ck: 2849 len: 280      1 Aab65251 Human PRO1131 (UNQ569) protein seq
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
191: SQSYS      (E)XFX(S)(Y)WXX(L)      LRPDS
      EFPYSYWTGL
1      AAB50959 ck: 2849 len: 280      1 Aab50959 Human PRO1131 protein, 3/2001
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
191: SQSYS      (E)XFX(S)(Y)WXX(L)      LRPDS
      EFPYSYWTGL
1      AAU19776 ck: 52 len: 394      1 Auu19776 Human novel extracellular matrix p
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
304: FLLEPF      (L)XFX(H)(F)WXX(F)      NALTL
      LFFGHFWQLF

```

```

1      ABB48814 ck: 2157 len: 510      1 Abb48814 Listeria monocytogenes protein
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
31: AQANK      (P)XFX(R)(Y)WXX(F)      PIGEL
      PNFDRIMADF
1      ABB49398 ck: 1607 len: 289      1 Abb49398 Listeria monocytogenes protein
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
175: HLNES      (E)XFX(S)(F)WXX(L)      RMTVT
      EDDFLSFNNNL
Databases searched:
EMBL, Release 7.0, Released on 28Apr2002, Formatted on 2May2002
Total finds: 158
Total length: 114,001,827
Total sequences: 766,495
CPU time: 14:23.28

```


*use accession # to watch
citation to sequence*

```

11AA_SEQUENCE 1.0
ID AAW32305 standard; Protein; 542 AA.
XX
AC AAW32305;
XX
DT 27-APR-1998 (first entry)
XX
DE Arabidopsis thaliana inorganic phosphate transporter 5.
XX
KM Columbia strain; inorganic phosphate transporter 5; IPT5;
XX accelerated phosphate uptake; tobacco plant.
XX
OS Arabidopsis thaliana.
XX
PN WO9735984-A1.
XX
PD 02-OCT-1997.
XX
PE 24-MAR-1997; 97MO-JP00975.
XX
PR 25-MAR-1996; 96JP-0094790.
XX
PA (MITS-) MITSUI PLANT BIOTECHNOLOGY RES INST.
XX
PI Mitsuoka N, Okumura S, Shibata D, Shirano Y;
XX
DR MPI: 1997-489647/45.
XX
DR N-PSDB; AAT91326.
XX
PT DNA encoding a plant phosphate transporter protein - useful for
XX producing e.g. tobacco plants with increased phosphate uptake and
XX accelerated growth
XX
PS Claim 2; Pages 80-83; 96pp; Japanese.
XX
CC The present sequence is Arabidopsis thaliana (columbia strain)
XX inorganic phosphate transporter 5 (IPT5), useful for accelerating
XX the uptake of phosphate by plants, e.g. tobacco plants.
XX
SO Sequence 542 AA;

AAW32305 Length: 542 May 30, 2002 09:25 Type: P Check: 992 ..

1 MARKGREVLN ALDAAKTQMY HPTAIVIAM GFETDAYDLF SISLVTKLIG
51 RIYHVDSSK KPGLTPNVA AAVNGVAFCG TLAGLFFGW LGDKLGRKKV
101 YGTTLMWVL CSLGSLFSG HSANGVMATL CEFRLFGLFG IGGDYPLSAT
151 IMSEYANKKT RGAFLAANFA MCGFGLIAG IVSLIVSTF DHAFAKPTYE
201 VDVGVSTVPQ ADYVNRIVLM FGAIFALLTY YMRKMPETA RYVALVAKND
251 KQASDMSKV LQVDLIAEEB AOSNSNSNP NTFGLFTRE FARRHGLHL
301 GTTTFELLD IAYSSNLFQ KDIYTAIGMI PAETMAIHI EPTVSKAQT
351 LIALCGTVPQ YWFTVAFIDI LGRFPIQMG FIFMTIMEA LAIPYDHRH
401 RNRRIFFLIM YSLTMEFANF GPNATTEYVP AEIIPARLRS TCHGISAASG
451 KAGAIWGAFC FLYAASSDS EKTDAGYPPG IGVNSLML ACVNFGLIVF
501 TLVPESSGK SLEISREDE EOSGDDIYVE MTVANSRKY PV

11AA_SEQUENCE 1.0
ID AAW32302 standard; Protein; 534 AA.
XX
AC AAW32302;
XX
DT 27-APR-1998 (first entry)
XX
DE Arabidopsis thaliana inorganic phosphate transporter 4.

```

```

XX
KM Columbia strain; inorganic phosphate transporter 4; IPT4;
XX accelerated phosphate uptake; tobacco plant.
XX
OS Arabidopsis thaliana.
XX
PN WO9735984-A1.
XX
PD 02-OCT-1997.
XX
PE 24-MAR-1997; 97MO-JP00975.
XX
PR 25-MAR-1996; 96JP-0094790.
XX
PA (MITS-) MITSUI PLANT BIOTECHNOLOGY RES INST.
XX
PI Mitsuoka N, Okumura S, Shibata D, Shirano Y;
XX
DR MPI: 1997-489647/45.
XX
DR N-PSDB; AAT91322, AAT91325.
XX
PT DNA encoding a plant phosphate transporter protein - useful for
XX producing e.g. tobacco plants with increased phosphate uptake and
XX accelerated growth
XX
PS Claim 2; Pages 72-74; 96pp; Japanese.
XX
CC The present sequence is Arabidopsis thaliana (columbia strain)
XX inorganic phosphate transporter 4 (IPT4), useful for accelerating
XX the uptake of phosphate by plants, e.g. tobacco plants.
XX
SO Sequence 534 AA;

AAW32302 Length: 534 May 30, 2002 09:25 Type: P Check: 605 ..

1 MAREOLOVLN ALDVAKTQMY HPTAIIAGM GFETDAYDLF CISLVTKLIG
51 RIYHVEGAQ KPGLTPNVA AAVNGVAFCG TLAGLFFGW LGDKLGRKKV
101 YGTTLMWVL CSLGSLFSG HEKAVMATL CEFRLFGLFG IGGDYPLSAT
151 IMSEYANKKT RGAFLVSAVFA MCGFGLIAG IFAIIISAF EAKPPSPAYA
201 DDALGSTIPQ ADLVNRILM AGAIPAMTY YSRKMPETA RYVALVAKDA
251 KQASDMSKV LQYEIEPEQ KLEISKEKS KAFGLFSKEF MSRGLHLIG
301 TTSWFELDI AFYSQNLFOK DIFSAIGWIP PAQSNMAIOE VFRIARAQTL
351 IALCSTVPQ WFTVAFIDVI GRFAIOMGF FFMVFMFAL AIPYHMTHK
401 ENRIGFVIMY SLTFEPANFG PNATTEYVPA EIPPARFST CHGISAASG
451 LGANVGAFGF LYLQNDPKD KTDAGYPPGI GVRNSLIVLG VVNFGLILFT
501 FLVPESSGKS LEEMSGENED NENSNDSTRT VPIV

11AA_SEQUENCE 1.0
ID AAW32303 standard; Protein; 521 AA.
XX
AC AAW32303;
XX
DT 27-APR-1998 (first entry)
XX
DE Arabidopsis thaliana inorganic phosphate transporter 3.
XX
KM Columbia strain; inorganic phosphate transporter 3; IPT3;
XX accelerated phosphate uptake; tobacco plant.
XX
OS Arabidopsis thaliana.
XX
PN WO9735984-A1.

```

PD 02-OCT-1997.
 XX
 PF 24-MAR-1997; 97WO-JP00975.
 XX
 PR 25-MAR-1996; 96JP-0094790.
 XX
 PA (MITS-) MITSUI PLANT BIOTECHNOLOGY RES INST.
 XX
 PI Mitsuoka N, Okumura S, Shibata D, Shitano Y;
 XX
 DR MPI: 1997-489647/45.
 DR N-PSDB; AAT91324.
 XX
 PT DNA encoding a plant phosphate transporter protein - useful for
 PT producing e.g. tobacco plants with increased phosphate uptake and
 PT accelerated growth
 PS
 PS Claim 2; Pages 75-77; 96pp; Japanese.
 CC The present sequence is Arabidopsis thaliana (columbia strain)
 CC inorganic phosphate transporter 3 (IPT3), useful for accelerating
 CC the uptake of phosphate by plants, e.g. tobacco plants.
 CC
 SQ Sequence 521 AA;

AAW32303 Length: 521 May 30, 2002 09:25 Type: P Check: 4821 ..

1 MADQQLGVLK ALDVAKTQLY HFTAIVIAGM GFFTDAYDLF CVSLVTKLLG
 51 RIYFNPESA KPGSLPPHYA AAVNGVALCG TLAGQLFFGW LGDKLGRKKV
 101 YGTLIMML CSVASGLSIG NSAKGVMTL CFFRFWLGFG IGDYPLSAT
 151 IMSEYANKKT RGAFLAAVFA MOGVGILAG FVALAVSSIF DKKFPAPTYA
 201 QDRFLSTPPQ ADYIWRIVM FGALPALTY YMRKMPETA RYTALVAKNI
 251 KQATADMSKY LQTDLEER VEDVYKDPK NYGLFSKEFL RHGHLHLCG
 301 TSTWFLDIA FYSQNLFOKD IFSALGMIPK AATMAIHEV FRIARAQTLI
 351 ALGSTVPGW FTVAFIDTIG RFAIQLNGF MMTVFMEALA FPNHMLP
 401 NRIGFVWMS LTFEFANFGP NATTFIVPAE IFFARLRSTC HGISAATGA
 451 GAIVGAFGL YAAPOQDKT TDAGYPPGIG VKNSLIMLV INFGMLTFE
 501 LVPEPKGKSL EELSGEAEVD K

!!AA_SEQUENCE 1.0
 ID AAW32304 standard; Protein; 524 AA.
 AC
 AC AAW32304;

27-APR-1998 (first entry)

DE Arabidopsis thaliana inorganic phosphate transporter 2.
 XX
 KW Columbia strain; inorganic phosphate transporter 2; IPT2;
 KW accelerated phosphate uptake; tobacco plant.
 OS
 OS Arabidopsis thaliana.

PN WO9735984-A1.

PD 02-OCT-1997.

PF 24-MAR-1997; 97WO-JP00975.

PR 25-MAR-1996; 96JP-0094790.

PA (MITS-) MITSUI PLANT BIOTECHNOLOGY RES INST.

PI Mitsuoka N, Okumura S, Shibata D, Shitano Y;
 XX
 DR MPI: 1997-489647/45.
 DR N-PSDB; AAT91324.
 XX
 PT DNA encoding a plant phosphate transporter protein - useful for
 PT producing e.g. tobacco plants with increased phosphate uptake and
 PT accelerated growth
 PS
 PS Claim 2; Pages 77-80; 96pp; Japanese.
 CC The present sequence is Arabidopsis thaliana (columbia strain)
 CC inorganic phosphate transporter 2 (IPT2), useful for accelerating
 CC the uptake of phosphate by plants, e.g. tobacco plants.
 CC
 SQ Sequence 524 AA;

AAW32304 Length: 524 May 30, 2002 09:25 Type: P Check: 8027 ..

1 MADQQLGVLK ALDVAKTQLY HFTAIVIAGM GFFTDAYDLF CVSLVTKLLG
 51 RIYFNPESA KPGSLPPHYA AAVNGVALCG TLAGQLFFGW LGDKLGRKKV
 101 YGTLIMML CSVASGLSFG NEAKGVMTL CFFRFWLGFG IGDYPLSAT
 151 IMSEYANKKT RGAFLAAVFA MOGVGILAG FVALAVSSIF DKKFPAPTYA
 201 VNRLSTPPQ VDIYRIIVM FGALPALTY YMRKMPETA RYTALVAKNI
 251 KQATADMSKY LQTDLEER VEDVYKDPK NYGLFSKEFL RHGHLHLCG
 301 TSTWFLDIA FYSQNLFOKD IFSALGMIPK AATMAIHEV FRIARAQTLI
 351 ALGSTVPGW FTVAFIDTIG RFAIQLNGF MMTVFMEALA FPNHMLP
 401 NRIGFVWMS LTFEFANFGP NATTFIVPAE IFFARLRSTC HGISAATGA
 451 GAIVGAFGL YAAONODKAK VDAGYPPGIG VKNSLIVLV INFGMLTFE
 501 LVPEPKGKSL EELSGEAEVS HDEK

!!AA_SEQUENCE 1.0
 ID AAW32301 standard; Protein; 524 AA.
 AC
 AC AAW32301;

27-APR-1998 (first entry)

DE Arabidopsis thaliana inorganic phosphate transporter 1.
 XX
 KW Columbia strain; inorganic phosphate transporter 1; IPT1;
 KW accelerated phosphate uptake; tobacco plant.
 OS
 OS Arabidopsis thaliana.

PN WO9735984-A1.

PD 02-OCT-1997.

PF 24-MAR-1997; 97WO-JP00975.

PR 25-MAR-1996; 96JP-0094790.

PA (MITS-) MITSUI PLANT BIOTECHNOLOGY RES INST.

PI Mitsuoka N, Okumura S, Shibata D, Shitano Y;

DR MPI: 1997-489647/45.

DR N-PSDB; AAT91321, AAT91323.

PT DNA encoding a plant phosphate transporter protein - useful for
 PT producing e.g. tobacco plants with increased phosphate uptake and
 PT accelerated growth

```

XX CC Claim 2: Pages 69-71; 96pp; Japanese.
XX CC
XX CC The present sequence is Arabidopsis thaliana (columbia strain)
CC CC inorganic phosphate transporter 1 (IPT1), useful for accelerating
CC CC the uptake of phosphate by plants, e.g. tobacco plants.
XX CC
SO Sequence 524 AA:

AAW32301 Length: 524 May 30, 2002 09:25 Type: P Check: 8746 ..
1 MAEQQLGVLK ALDVAKTQLY HFTAIIVAGM GFETDAYDLF CYSIVTKLIG
51 RIYFNPESA KPGSLPPHYA AAVNGVALCG TLSGQLEFGW LGDKLGRKKV
101 YGLTVMVIL CSVASGLSPG HEAKGVMTLL CFFRFLGFG IGGDYPISAT
151 IMSEYANKKT KGAFIAVFA MGVGILAGG FVALAVSSIF DKRPAPTYA
201 VNRLSTPQ VDIYIRIIV FGLPALATY YWRMKMPETA RYVALVAKNI
251 KQATADMSKV LQTDIELEER VEDDYKDPKQ NYGLFSKEFL RRHGLHLIGT
301 TSTWFLDLIA FYSQNLFOKD IFSAGIWIPI AATMNATHEV FRIARAQTLI
351 ALCTVPGYW FTVAFIDTIG RPKIQLNGPF MMTVEMFAIA FPNHWIKPE
401 NRICEFVWYS LTFEFANFGP NATTFIVPAE IFPARLRSTC HGISAAGAKA
451 GAIYGAFGL YAAQSDOKAR VDAGYPPGIG VKNSLIMLGV LNFGLMLTF
501 LVPEPKGKSL EELSGEAEVS HDEK

11AA_SEQUENCE 1.0
ID AAW30486 standard; Protein: 135 AA.
XX AC
XX AAW30486:
XX DT
XX DT 14-APR-1998 (first entry)
XX DE
XX DE Flea saliva protein fspn (Pfspn6-135).
XX KM
XX KM Flea saliva protein; fspn; allergic dermatitis; allergy;
XX KM therapy; diagnosis; antibody; Pfspn6-135.
XX OS
XX OS Ctenocephalides felis.
XX PN
XX PN WO9737676-A1.
XX PD
XX PD 16-OCT-1997.
XX PF
XX PF 10-APR-1997; 97WO-US05959.
XX PR
XX PR 10-APR-1996; 9605-0630822.
XX PA
XX PA (HESK-) HESKA CORP.
XX PI
XX PI Hunter SW, Sim G, Weber ER;
XX DR
XX DR WPI; 1997-512409/47.
XX DR N-PSDB; AAT92826.
XX PT
XX PT New flea saliva proteins - useful for treating allergic dermatitis
XX PT and as diagnostic reagents
XX PS
XX PS Claim 2; Page 156; 179pp; English.
XX CC
XX CC This polypeptide comprises a non-full-length flea saliva protein
CC CC (FSP), denoted Pfspn6-135, that can be used to treat allergic
CC CC dermatitis. Its amino acid sequence was deduced from nucleic
CC CC acid nfpn6-406 (see AAT92826). Claimed PSP polypeptides (see also
CC CC AAW30480, AAW30483, AAW30487 and AAW30488-91) can be expressed in host
CC CC cells. The proteins, or their fragments or mimetopes, are used in

```

```

CC CC claimed methods for treating allergic dermatitis in animals, to
CC CC determine if an animal is susceptible to, or has, allergic
CC CC dermatitis, and to desensitise a host animal to allergic
CC CC dermatitis, as well as to monitor progress or effects of treatment.
CC CC Also contemplated is the in vivo expression of FSPs. FSPs can also
CC CC be used to raise antibodies useful as immunoassay reagents and for
CC CC passive immunisation.
XX CC
SO Sequence 135 AA:

AAW30486 Length: 135 May 30, 2002 09:25 Type: P Check: 8893 ..
1 MYKGPDEAC NYAGGPQLTT LQEKDSVLE DGKTEAYELG KILDRVYKKQ
51 LKVDKWDATK TYMAVSTKAM RTEKALIVG AGLENNPAKA KGNNTQOQLD
101 STHEDPMPGF SREWNPOCP AYPRALSLOL QIKIK

11AA_SEQUENCE 1.0
ID AAW36137 standard; Protein: 534 AA.
XX AC
XX AAW36137:
XX DT
XX DT 25-MAR-1998 (first entry)
XX DE
XX DE A. thaliana inorganic phosphate transporter protein IPT4.
XX KM
XX KM Promoter; inorganic phosphate transporter gene; IPT4; expression;
XX KM regulation; transgenic plant; expression vector.
XX OS
XX OS Arabidopsis thaliana.
XX PN
XX PN JP09252782-A.
XX PD
XX PD 30-SEP-1997.
XX PF
XX PF 25-MAR-1996; 96JP-0094856.
XX PR
XX PR 25-MAR-1996; 96JP-0094856.
XX PA
XX PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
XX DR
XX DR WPI; 1997-530154/49.
XX DR N-PSDB; AAV01455.
XX PT
XX PT A promoter for a phosphate transporter gene - useful for regulating
XX PT heterologous gene expression in transgenic plants
XX PS
XX PS Example 2; Page 6-9; 12pp; Japanese.
XX CC
XX CC This is the amino acid sequence of the Arabidopsis thaliana inorganic
CC CC phosphate transporter gene IPT4. The promoter region (bases 1-1875; see
CC CC also AAV01454) of the encoding gene can be used to regulate the level of
CC CC expression of a heterologous gene in a plant by introducing, into a
CC CC plant, a vector having the heterologous gene linked downstream region
CC CC from the promoter and then regulating extracellular phosphate levels.
XX CC
SO Sequence 534 AA:

AAW36137 Length: 534 May 30, 2002 09:25 Type: P Check: 605 ..
1 MAREQLOVLN ALDVAKTQWY HFTAIIVAGM GFETDAYDLF CYSIVTKLIG
51 RIYHVEGAQ KPGTLPPHYA AAVNGVAFGC TLAGQLEFGW LGDKLGRKKV
101 YGMLTMVVL CSIASGLSPG HEPRKAVWATL CFFRFLGFG IGGDYPISAT
151 IMSEYANKKT KGAFVSAVFA MGVGIMAGG IFATISSAF EAKPPSPAYA
201 DDALGSTIPQ ADLYWRITLM AGAIPAAWY YRSKMPETA RYVALVAKDA
251 KQATADMSKV LQVEIEPEQO KLEIEISERS KAFGLFSKEF MSRIGLHLIG

```

301 TTSTWFLDDI AFYSONLPFOK DIFSAICWIP PAOSMAIOE VFRIARAQTL
 351 IALCSTVPGT WETVAFIDVI GRFAIQMGF FEMTFEMFAL AIPYNWHTK
 401 ENRIGEVIM SLTFFANFG PNATTFVPA EIFPARRST CHGISASGK
 451 LGAMVGAFCF LYLQNPDKD KTDAGYPRGI GVRNSLVLG VVNFGLILFT
 501 FLVPEKSGKS LEMSGENED NENSNDSTRT VPIV

11AA_SEQUENCE 1.0

ID AAW82377 standard; Protein; 135 AA.

AC AAW82377;

DT 26-APR-1999 (first entry)

DE Flea saliva protein Pfspn6-135.

KW Flea saliva protein; Pfspn6-135; allergic dermatitis; allergen;
 allergy; therapy; diagnosis; vaccine; ectoparasite.

OS Ctenocephalides sp.

PN WO9845408-A2.

PD 15-OCT-1998.

PF 15-OCT-1997; 97WO-US18669.

PR 10-APR-1997; 97WO-US05959.

PA (HESK-) HESKA CORP.

PI Weber ER;

DR WPI; 1998-594480/50.

DR N-PSDB; AAV73417.

PT New isolated ectoparasite saliva genes - used to develop products
 for the diagnosis, prevention, treatment and determining
 susceptibility to allergic dermatitis

PS Example 10; Page 125; 172pp; English.

CC This is flea saliva protein Pfspn6-135, encoded by a cDNA clone,
 nfp6-405 (see AAV73417), that was isolated from a flea salivary
 gland cDNA library. Pfspn6-135 shows no significant homology to
 known amino acid sequences. The invention relates to novel methods
 for isolating ectoparasite saliva proteins (ESPs), including flea
 saliva proteins. It provides ESps (claimed, see AAW82382-93), nucleic
 acids encoding them, methods for their recombinant production,
 therapeutic compositions for treating allergic dermatitis that
 comprise at least one ESP, as well as assay kits for testing if an
 animal has, or is susceptible to, allergic dermatitis, and a method
 of desensitising a host animal to allergic dermatitis using ESps.
 CC The ESps can also be used for the production of antibodies useful
 in diagnosis or in vaccines for passive immunisation against
 allergic dermatitis.

CC Sequence 135 AA;

AAW82377 Length: 135 May 30, 2002 09:25 Type: P Check: 8893 ..

1 MVKGPDEHC NYAGPQULT LQEKDSVLTG DGTAEVLEIG KILDVYKRO

51 LKVDKMDATK TYMAVSTKAM RTKEALIVG AGLENNPAPA KGNWTOQOOLD

101 SHFEDAMPGF SRFWNPQCCP AYFALSLQON QKTKK

11AA_SEQUENCE 1.0

ID AAW82382 standard; Protein; 375 AA.

XX

AC AAW82382;

DT 26-APR-1999 (first entry)

DE Flea saliva protein Pfspn6-375.

KW Flea saliva protein; Pfspn6-375; allergic dermatitis; allergen;
 allergy; therapy; diagnosis; vaccine; ectoparasite.

OS Ctenocephalides sp.

PN Key Location/Qualifiers

FT Peptide 1..19 /note="signal peptide"

FT Protein 20..375 /note="mature protein"

PN WO9845408-A2.

PD 15-OCT-1998.

PF 15-OCT-1997; 97WO-US18669.

PR 10-APR-1997; 97WO-US05959.

PA (HESK-) HESKA CORP.

PI Weber ER;

DR WPI; 1998-594480/50.

DR N-PSDB; AAV73430 AND AAV73432.

PT New isolated ectoparasite saliva genes - used to develop products
 for the diagnosis, prevention, treatment and determining
 susceptibility to allergic dermatitis

PS Claim 2; Page 133-134; 172pp; English.

CC This polypeptide comprises a new full-length flea saliva protein,
 CC termed Pfspn6-375. Its amino acid sequence was deduced from a
 CC cDNA clone (see AAV73430) isolated from a whole flea cDNA library,
 CC and shows no significant homology to known amino acid sequences.
 CC The mature fspn6-356 protein (see AAW82384) has been expressed
 CC in recombinant Escherichia coli BL-21 cells. The invention is
 CC directed to methods for isolating ectoparasite saliva proteins
 CC (ESPs). It provides ESps (claimed, see AAW82382-93), nucleic acid
 CC molecules encoding them, methods for their recombinant production,
 CC therapeutic compositions for treating allergic dermatitis that
 CC comprise at least one ESP, as well as assay kits for testing if an
 CC animal has, or is susceptible to, allergic dermatitis, and a method
 CC of desensitising a host animal to allergic dermatitis using ESps.
 CC The ESps can also be used for the production of antibodies useful
 CC in diagnosis or in vaccines for passive immunisation against
 CC allergic dermatitis.

CC Sequence 375 AA;

AAW82382 Length: 375 May 30, 2002 09:25 Type: P Check: 6654 ..

1 MFATILTGIL ALTLSECEB LKFVPMWKG PDHEACWAG GPOLTLQEK

51 DSVLTEDGKT EAVEIGKLLD KYKKQOLKVD KWDATKTYMA VSTKAMRTKE

101 ALIVGAGLE NNPBARKGNW TQOQLDSTHF DAMPGFSRFW NPOQCPAYFR

151 ALSIQNQKIK KLEEKYQITI KEVTAKPFSI DGTAGQIIMI AYTFEKKMQ

201 QGRKEVEGIN YATMOKLKEP SSEFVLIALT STDQMRKLAG GLILKDLFND

251 IDELTCKDHAQ PHAFGIRKKN NNIFFVPOAI LAAQMAVEMP EGTKLDDPI

301 TASNFTDDQ SYVITIELYOD KKKMNVOLQY KKKNSGMLP IKVQGCNSPM

351 CPYDTLKSL NKYIIDARH KOACK

11AA_SEQUENCE 1.0

ID AAM82384 standard; Protein; 356 AA.

AC AAM82384;

DT 26-APR-1999 (first entry)

DE Flea saliva protein Pfspn6-356.

KW Flea saliva protein; Pfspn6-356; allergic dermatitis; allergen;
allergy; therapy; diagnosis; vaccine; ectoparasite.

OS Ctenocephalides sp.

PN MO9845408-A2.

PD 15-OCT-1998.

PF 15-OCT-1997; 97WO-US18669.

PR 10-APR-1997; 97WO-US05959.

PA (HESK-) HESKA CORP.

PI Weber ER;

DR WPI: 1998-594480/50.

DR N-PSDB; AAV73434.

PT New isolated ectoparasite saliva genes - used to develop products
for the diagnosis, prevention, treatment and determining
susceptibility to allergic dermatitis

PS Claim 2; Page 137; 172pp; English.

This polypeptide comprises a new, claimed 40.5 kDa flea saliva protein, termed Pfspn6-356, consisting of the mature polypeptide region of Pfspn6-375 (see AAM82382). Pfspn6-356 has been expressed in recombinant *Escherichia coli* BL-21 cells. The invention is directed to methods for isolating ectoparasite saliva proteins (ESPs). It provides ESps (claimed, see AAM82382-93), nucleic acid molecules encoding them, methods for their recombinant production, therapeutic compositions for treating allergic dermatitis that comprise at least one ESP, as well as assay kits for testing if an animal has, or is susceptible to, allergic dermatitis, and a method of desensitising a host animal to allergic dermatitis using ESps. The ESps can also be used for the production of antibodies useful in diagnosis or in vaccines for passive immunisation against allergic dermatitis.

SO Sequence 356 AA;

AAM82384 Length: 356 May 30, 2002 09:25 Type: P Check: 2624 ..

1 ELKFVFWVK GPDHACNYA GPQLTTIQE KDSVLTEDGK TEAYELGKLL

51 DKYKQOLAV DKWDATKTYW AVSTRAMRTK EALILYGAGL ENNPAAKAGN

101 WTOOQLDSTH FDMAPGFSRF WNPQCCPAYF RALSLQNOKI KLLKXYQTT

151 IREVTAKKFS IDGTAKOHIM IAYEFKRMK QQGRKEVEGI NRTMOKLKE

201 FSSEFVLIAL TSTDOMRKLA GGLILKDLFN DDELTKDHA OPHAPGICIN

251 KNMIFVVPQA ILAQMAYFM PEGTKLRDQF ITASNFYDDO QSYVIELYO

301 DKKMNVOLO YKNNKSGWL PIYOGCNSP MCPYDTLKSL LNKYIIDARH

351 HKQACK

11AA_SEQUENCE 1.0

ID AAM82385 standard; Protein; 355 AA.

AC AAM82385;

DT 26-APR-1999 (first entry)

DE Flea saliva protein Pfspn6-357.

KW Flea saliva protein; Pfspn6-357; allergic dermatitis; allergen;
allergy; therapy; diagnosis; vaccine; ectoparasite.

OS Ctenocephalides sp.

PN MO9845408-A2.

PD 15-OCT-1998.

PF 15-OCT-1997; 97WO-US18669.

PR 10-APR-1997; 97WO-US05959.

PA (HESK-) HESKA CORP.

PI Weber ER;

DR WPI: 1998-594480/50.

DR N-PSDB; AAV73436.

PT New isolated ectoparasite saliva genes - used to develop products
for the diagnosis, prevention, treatment and determining
susceptibility to allergic dermatitis

PS Example 10; Page 137; 172pp; English.

This polypeptide comprises the new, claimed 40.5 kDa flea saliva protein Pfspn6-356 (see AAM82384) plus an additional N-terminal Met residue. It was expressed in *Escherichia coli* BL-21 cells transformed with a vector carrying a PCR amplification product (see AAV73436) termed Pfspn6-1071. The invention is directed to products and methods for isolating ectoparasite saliva proteins (ESPs). It provides ESps (claimed, see AAM82382-93), nucleic acid molecules encoding them, methods for their recombinant production, therapeutic compositions for treating allergic dermatitis that comprise at least one ESP, as well as assay kits for testing if an animal has, or is susceptible to, allergic dermatitis, and a method of desensitising a host animal to allergic dermatitis using ESps. The ESps can also be used for the production of antibodies useful in diagnosis or in vaccines for passive immunisation against allergic dermatitis.

SO Sequence 355 AA;

AAM82385 Length: 355 May 30, 2002 09:25 Type: P Check: 309 ..

1 MELKFVFKG PDHACNYAG GPQLTTIQEK DSVLTEDGKT EAYELGKLLD

51 KYVKKQOLKVD KWDATKTYWA VSTRAMRTKE AALILYGAGLE NNPAAKAGNW

101 TQOQLDSTHF DAMPGFSRFV NPQCCPAYFR ALSLQNOKIK KLLKXYQTTT

151 KEVTAKKPSSI DGTAKOHIMI AYETFRMKO QGRKEVEGIN TATMOKLKEE

201 SSEFVLIALT STDOMRKLAG GLLLDLPND IDELTCKDHAQ PHAPGICIKN

251 MNIFVVPQAI LAAQMAVFM PEGTKLRDQPI TASNFYDDOQ SYVIELYOD

301 KKNMNVOLOY KNNKSGWLP IKVGCNSPM CPYDTLKSL NKYIIDARH

351 KOACK

11AA_SEQUENCE 1.0

ID AAM82322 standard; Peptide; 19 AA.

XX

AC AAW82322;
DT 22-FEB-1999 (first entry)
XX
DE p53 homologue TIP 12/1 peptide.
XX
KW p53; mdm2; inhibitor; therapy; activator; treatment; cancer; medicament.
XX
OS Synthetic.
XX
PN WO9847919-A1.
XX
PD 29-OCT-1998.
XX
PF 20-APR-1998; 98WO-GB01140.
XX
PR 22-APR-1997; 97GB-0008089.
XX
PA (UYDU-) UNIV DUNDEE.
XX
PI Lane DP;
XX
DR WPI; 1998-609975/51.
XX
PT New substance with a mdm2 binding domain and coupling partner -
PT useful for stabilising in cells without an efficient mdm2-mediated
PT degradation pathway
XX
PS Disclosure; Fig 1; 52pp; English.
XX
CC This sequence is a peptide homologue of a region of p53 which binds
CC to mdm2. This peptide is used in the construction of a novel agent
CC capable of disrupting the binding of p53 and mdm2 or inhibiting the
CC production of mdm2 in a population of cells. This agent is also used in
CC the preparation of a therapeutic for activating p53, where the population
CC of cells do not overexpress mdm2. Inhibiting mdm2 production and/or
CC inhibiting the binding of mdm2 to p53 allows levels of p53 to increase by
CC reducing the clearance of p53 by mdm2, and can be used to activate p53
CC function. The agents for use in therapeutics for activating p53 can be
CC used for the treatment of cancer, viral conditions or other conditions
CC associated with non-functional p53.
XX
SQ Sequence 19 AA;

AAW82322 Length: 19 May 30, 2002 09:25 Type: P Check: 4400 ..

1 PPLSMRFRMD YWEGLENENG

!!AA_SEQUENCE 1.0
ID AAW82320 standard; Peptide; 19 AA.
XX
AC AAW82320;

DT 22-FEB-1999 (first entry)
XX

DE p53 homologue TIP 12/1 peptide.
XX

KW p53; mdm2; inhibitor; therapy; activator; treatment; cancer; medicament.
XX

OS Synthetic.
XX

PN WO9847525-A1.
XX

PD 29-OCT-1998.
XX

PF 20-APR-1998; 98WO-GB01144.
XX

PR 22-APR-1997; 97GB-0008092.
XX

PA (UYDU-) UNIV DUNDEE.
XX

PI Lane DP;
XX

DR WPI; 1998-609932/51.
XX
XX New agents which inhibit interaction of p53 and mdm2 - useful for
PT activating p53, e.g. for treating cancers; viral conditions or other
PT conditions associated with non functional p53 or mdm2
XX
XX Disclosure; Fig 1; 52pp; English.
XX
CC This sequence is a peptide homologue of a region of p53 which binds
CC to mdm2. This peptide is used in the construction of a novel agent
CC capable of disrupting the binding of p53 and mdm2 or inhibiting the
CC production of mdm2 in a population of cells. This agent is also used in
CC the preparation of a therapeutic for activating p53, where the population
CC of cells do not overexpress mdm2. Inhibiting mdm2 production and/or
CC inhibiting the binding of mdm2 to p53 allows levels of p53 to increase by
CC reducing the clearance of p53 by mdm2, and can be used to activate p53
CC function. The agents for use in therapeutics for activating p53 can be
CC used for the treatment of cancer, viral conditions or other conditions
CC associated with non-functional p53.
XX
SQ Sequence 19 AA;

AAW82320 Length: 19 May 30, 2002 09:25 Type: P Check: 4400 ..

1 PPLSMRFRMD YWEGLENENG

!!AA_SEQUENCE 1.0
ID AAW58852 standard; Protein; 130 AA.
XX

AC AAW58852;

DT 23-JUL-1998 (first entry)
XX

DE Human O289_1 secreted protein.
XX

KW Secreted protein; prevention; treatment; gene therapy.
XX

OS Homo sapiens.
XX

PN WO9801554-A2.
XX

PD 15-JAN-1998.
XX

PF 07-JUL-1997; 97WO-US11876.
XX

PR 09-JUL-1996; 96US-0677231.
XX

PA (GEMV) GENETICS INST INC.
XX

PI Bowman M, Evans C, Jacobs K, LaVallie ER, McCoy JM;
PI Merberg D, Racie LA, Spaulding V, Treacy M;

DR WPI; 1998-110230/10.
XX

DR N-PSDB; AAV11438, AAV11439.
XX

PT Secreted proteins and polynucleotides encoding them - useful to
PT prevent, treat and ameliorate medical conditions
XX

PS Disclosure; Page 71; 93pp; English.
XX

CC This sequence represents a fragment of a novel secreted protein derived
CC from clone O289_1 which was isolated from a human dendritic cell cDNA
CC library. The protein can be used to prevent, treat or ameliorate a
CC medical condition, while the polynucleotides can be used for gene
CC therapy.
XX

SQ Sequence 130 AA;

AAW58852 Length: 130 May 30, 2002 09:25 Type: P Check: 5772 ..

1 NENPGVCHSS ATGDPHCROP EPTSPGPLSW LFLSLFPFPPR LAVTHRETSQ

51 LGVODRDFRY FLSKERBDF YGSLPVGILL AHRVDHTKSC QGFPLSIPTP

101 SPRVRYTALI QALVISHLCH FSSFWIGLPA
11AA_SEQUENCE 1.0
ID AAM56025 standard; protein; 316 AA.
XX
AC AAM56025;
XX
DT 28-JUL-1998 (first entry)
XX
DE Phosphate starvation-induced protein kinase psr6.
XX
KW Phosphate starvation-induced protein kinase; psr6; psr1; glucosidase;
KM phosphate transporter; phosphate deprivation; Arabidopsis thaliana;
KW Brassica nigra; photosynthetic organism; phosphorus metabolism; growth;
KM reproduction; metabolic content; flowering; drought; cold tolerance.
XX
OS Brassica nigra.
XX
PN WO9805760-A2.
XX
PD 12-FEB-1998.
XX
PE 30-JUL-1997; 97MO-CA00532.
XX
PR 31-JUL-1996; 96US-0688988.
PR 31-JUL-1996; 96CA-2182421.
XX
PA (TOOH) UNIV QUEBENS KINGSTON.
XX
PI Lefebvre DD, Malboobi MA;
XX
DR MPI; 1998-159169/14.
DR N-PSDB; AAV26239.
XX
PT New isolated phosphate starvation-inducible genes - used to develop
PT plants or other photosynthetic organisms with altered phosphorus
PT metabolism, e.g. to alter growth or reproduction.
XX
PS Claim 62; Fig 21; 131pp; English.
XX
CC The present sequence represents a phosphate starvation-induced protein
CC kinase from a photosynthetic organism, in which transcription of the
CC DNA is induced by phosphate deficiency. The products and methods of
CC the present invention can be used to modify the phosphorus metabolism
CC in plants and other photosynthetic organisms, e.g. to alter growth,
CC reproduction, metabolic content, flowering, drought or cold tolerance
CC or nutritive value.
XX
SQ Sequence 316 AA;
AAM56025 Length: 316 May 30, 2002 09:25 Type: P Check: 9188 ..
1 IPGKTLNVMV LCSVASGLSF GDRKPSVMTT LCFRFLWGF GIGDYPLSA
51 TINSEYANKR TRGAFYSAYF AMQGFIMAG GIFAIISSA FFAKFPAPAY
101 AEDALASTVP QADFVWRIL MWGAIPAMT YYSRSKMPET ARTALVARD
151 AKQASDMSR VLQVEIEAQ EKVEISSNK SKAFSLFSKE FMKRHGIHL
201 GTTSTWFLD IAFYSQNLQ KDIFSAIGMI PPAOTMNAIQ EYFKIARAQT
251 LIALCSTVPQ YMFYVAFIDV IGRFAIQMG FFEFTVFMFA LALPYNHWTH
301 TRNNLNCWKL KKRLLD
11AA_SEQUENCE 1.0
ID AAM37220 standard; peptide; 12 AA.
XX
AC AAM37220;
XX
DT 20-JUL-1998 (first entry)

XX
DE MDM2 binding peptide unique phage insert sequence 1.
XX
KM MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KM tumour; diagnosis; binding; viral infection; phage insert.
XX
OS Homo sapiens.
XX
PN WO9801467-A2.
XX
PD 15-JAN-1998.
XX
PE 04-JUL-1997; 97MO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picklesley S;
XX
DR MPI; 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 8; Page 30; 45pp; English.
XX
CC This is a unique phage insert sequence of the MDM2 binding peptide
CC identified by phage display. The MDM2 binding peptides and their
CC derivatives are capable of binding to the human oncogenic protein MDM2.
CC These peptides can specifically inhibit or block the binding of MDM2 to
CC the human p53 protein, in vitro or in vivo. Inhibiting the interaction
CC between the p53 and MDM2 can induce growth arrest or apoptosis in tumour
CC cells comprising a wild-type p53 and non-elevated levels of MDM2. The
CC peptides may be used to identify molecules that bind to MDM2 and to
CC identify and design inhibitors of MDM2/p53 binding. They may also be
CC used to purify binding partners especially MDM2, diagnose disease by
CC measuring levels of MDM2 in blood of cancer and leukaemia patients and
CC for treatment or prevention of disease involving p53/MDM2 interactions,
CC especially tumours and viral infections. The peptides can be administered
CC nasally, rectally, orally or by injection. By interfering with MDM2/p53
CC interaction, the peptides can activate p53 function and accumulation in
CC normal cells. The peptides which mimic the MDM2 binding site in p53, have
CC a significantly greater blocking activity compared with wild-type p53.
XX
SQ Sequence 12 AA;
AAM37220 Length: 12 May 30, 2002 09:25 Type: P Check: 5978 ..
1 MPRPMYWEGL LN
11AA_SEQUENCE 1.0
ID AAM37224 standard; peptide; 15 AA.
XX
AC AAM37224;
XX
DT 20-JUL-1998 (first entry)
XX
DE MDM2 binding peptide unique phage insert sequence 5.
XX
KM MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KM tumour; diagnosis; binding; viral infection; phage insert.
XX
OS Homo sapiens.
XX
PN WO9801467-A2.
XX
PD 15-JAN-1998.
XX

```

PF 04-JUL-1997; 97WO-EP03549.
XX
XX 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS ) NOVARTIS AG.
XX
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Plicksley S;
XX
XX WPI; 1998-100996/09.
DR
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
XX infections and identifying binding agents
XX
XX Example 8; Page 30; 45pp; English.
XX
XX This is a unique phage insert sequence of the MDM2 binding peptide
CC identified by phage display. The MDM2 binding peptides and their
CC derivatives are capable of binding to the human oncogenic protein MDM2.
CC These peptides can specifically inhibit or block the binding of MDM2 to
CC the human p53 protein, in vitro or in vivo. Inhibiting the interaction
CC between the p53 and MDM2 can induce growth arrest or apoptosis in tumour
CC cells comprising a wild-type p53 and non-elevated levels of MDM2. The
CC peptides may be used to identify molecules that bind to MDM2 and to
CC identify and design inhibitors of MDM2/p53 binding. They may also be
CC used to purify binding partners especially MDM2, diagnose disease by
CC measuring levels of MDM2 in blood of cancer and leukaemia patients and
CC for treatment or prevention of disease involving p53/MDM2 interactions,
CC especially tumours and viral infections. The peptides can be administered
CC nasally, rectally, orally or by injection. By interfering with MDM2/p53
CC interaction, the peptides can activate p53 function and accumulation in
CC normal cells. The peptides which mimic the MDM2 binding site in p53, have
CC a significantly greater blocking activity compared with wild-type p53.
XX
XX Sequence 15 AA:
SQ
AAM37224 Length: 15 May 30, 2002 09:25 Type: P Check: 9428 ..
1 PRPALVFADY WETLY
!!AA_SEQUENCE 1.0
ID AAM37221 standard; peptide; 12 AA.
XX
XX AAM37221;
AC
XX
XX 20-JUL-1998 (first entry)
DT
XX
XX MDM2 binding peptide unique phage insert sequence 2.
DE
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX tumour; diagnosis; binding; viral infection; phage insert.
XX
XX Homo sapiens.
OS
XX
XX W09801467-A2.
PN
XX
XX 15-JAN-1998.
PI
XX
XX 04-JUL-1997; 97WO-EP03549.
PF
XX
XX 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS ) NOVARTIS AG.
XX
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Plicksley S;
XX
XX WPI; 1998-100996/09.
DR

```

```

XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
XX infections and identifying binding agents
XX
XX Example 8; Page 30; 45pp; English.
XX
XX This is a unique phage insert sequence of the MDM2 binding peptide
CC identified by phage display. The MDM2 binding peptides and their
CC derivatives are capable of binding to the human oncogenic protein MDM2.
CC These peptides can specifically inhibit or block the binding of MDM2 to
CC the human p53 protein, in vitro or in vivo. Inhibiting the interaction
CC between the p53 and MDM2 can induce growth arrest or apoptosis in tumour
CC cells comprising a wild-type p53 and non-elevated levels of MDM2. The
CC peptides may be used to identify molecules that bind to MDM2 and to
CC identify and design inhibitors of MDM2/p53 binding. They may also be
CC used to purify binding partners especially MDM2, diagnose disease by
CC measuring levels of MDM2 in blood of cancer and leukaemia patients and
CC for treatment or prevention of disease involving p53/MDM2 interactions,
CC especially tumours and viral infections. The peptides can be administered
CC nasally, rectally, orally or by injection. By interfering with MDM2/p53
CC interaction, the peptides can activate p53 function and accumulation in
CC normal cells. The peptides which mimic the MDM2 binding site in p53, have
CC a significantly greater blocking activity compared with wild-type p53.
XX
XX Sequence 12 AA:
SQ
AAM37221 Length: 12 May 30, 2002 09:25 Type: P Check: 6151 ..
1 VQNFIDYWTQ QF
!!AA_SEQUENCE 1.0
ID AAM37222 standard; peptide; 12 AA.
XX
XX AAM37222;
AC
XX
XX 20-JUL-1998 (first entry)
DT
XX
XX MDM2 binding peptide unique phage insert sequence 3.
DE
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX tumour; diagnosis; binding; viral infection; phage insert.
XX
XX Homo sapiens.
OS
XX
XX W09801467-A2.
PN
XX
XX 15-JAN-1998.
PI
XX
XX 04-JUL-1997; 97WO-EP03549.
PF
XX
XX 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS ) NOVARTIS AG.
XX
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Plicksley S;
XX
XX WPI; 1998-100996/09.
DR
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
XX infections and identifying binding agents
XX
XX Example 8; Page 30; 45pp; English.
XX
XX This is a unique phage insert sequence of the MDM2 binding peptide
CC identified by phage display. The MDM2 binding peptides and their
CC derivatives are capable of binding to the human oncogenic protein MDM2.
CC These peptides can specifically inhibit or block the binding of MDM2 to
CC the human p53 protein, in vitro or in vivo. Inhibiting the interaction

```


CC between the p53 and MDM2 can induce growth arrest or apoptosis in tumour
 CC cells comprising a wild-type p53 and non-elevated levels of MDM2. The
 CC peptides may be used to identify molecules that bind to MDM2 and to
 CC identify and design inhibitors of MDM2/p53 binding. They may also be
 CC used to purify binding partners especially MDM2, diagnose disease by
 CC measuring levels of MDM2 in blood of cancer and leukaemia patients and
 CC for treatment or prevention of disease involving p53/MDM2 interactions,
 CC especially tumours and viral infections. The peptides can be administered
 CC nasally, rectally, orally or by injection. By interfering with MDM2/p53
 CC interaction, the peptides can activate p53 function and accumulation in
 CC normal cells. The peptides which mimic the MDM2 binding site in p53, have
 CC a significantly greater blocking activity compared with wild-type p53.

XX Sequence 12 AA:

AAW37222 Length: 12 May 30, 2002 09:25 Type: P Check: 5993 ..

1 TGPFTHYWA TF

11AA_SEQUENCE 1.0
 ID AAW37223 standard; peptide: 15 AA.

AC AAW37223;

DT 20-JUL-1998 (first entry)

XX MDM2 binding peptide unique phage insert sequence 4.

XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
 KM tumour; diagnosis; binding; viral infection; phage insert.

OS Homo sapiens.

XX MO9801467-A2.

PD 15-JAN-1998.

PF 04-JUL-1997; 97WO-EP03549.

PR 07-APR-1997; 97GB-0007041.

PT 05-JUL-1996; 96GB-0014197.

XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.

PA (NOVS) NOVARTIS AG.

PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
 PI Hochkeppel H, Lane D, Pickersley S;

DR WPI: 1998-100996/09.

XX Compounds binding to MDM2 protein and inhibit its interaction with
 PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
 PT infections and identifying binding agents

XX Example 8; Page 30; 45pp; English.

CC This is a unique phage insert sequence of the MDM2 binding peptide
 CC identified by phage display. The MDM2 binding peptides and their
 CC derivatives are capable of binding to the human oncogenic protein MDM2.
 CC These peptides can specifically inhibit or block the binding of MDM2 to
 CC the human p53 protein, in vitro or in vivo. Inhibiting the interaction
 CC between the p53 and MDM2 can induce growth arrest or apoptosis in tumour
 CC cells comprising a wild-type p53 and non-elevated levels of MDM2. The
 CC peptides may be used to identify molecules that bind to MDM2 and to
 CC identify and design inhibitors of MDM2/p53 binding. They may also be
 CC used to purify binding partners especially MDM2, diagnose disease by
 CC measuring levels of MDM2 in blood of cancer and leukaemia patients and
 CC for treatment or prevention of disease involving p53/MDM2 interactions,
 CC especially tumours and viral infections. The peptides can be administered
 CC nasally, rectally, orally or by injection. By interfering with MDM2/p53
 CC interaction, the peptides can activate p53 function and accumulation in
 CC normal cells. The peptides which mimic the MDM2 binding site in p53, have
 CC a significantly greater blocking activity compared with wild-type p53.

XX Sequence 15 AA:

AAW37223 Length: 15 May 30, 2002 09:25 Type: P Check: 9093 ..

1 IDRAPTFRDH WFALV

11AA_SEQUENCE 1.0
 ID AAW37225 standard; peptide: 15 AA.

AC AAW37225;

DT 20-JUL-1998 (first entry)

XX MDM2 binding peptide unique phage insert sequence 6.

XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
 KM tumour; diagnosis; binding; viral infection; phage insert.

OS Homo sapiens.

XX MO9801467-A2.

PD 15-JAN-1998.

PF 04-JUL-1997; 97WO-EP03549.

PR 07-APR-1997; 97GB-0007041.

PT 05-JUL-1996; 96GB-0014197.

XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.

PA (NOVS) NOVARTIS AG.

PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
 PI Hochkeppel H, Lane D, Pickersley S;

DR WPI: 1998-100996/09.

XX Compounds binding to MDM2 protein and inhibit its interaction with
 PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
 PT infections and identifying binding agents

XX Example 8; Page 30; 45pp; English.

CC This is a unique phage insert sequence of the MDM2 binding peptide
 CC identified by phage display. The MDM2 binding peptides and their
 CC derivatives are capable of binding to the human oncogenic protein MDM2.
 CC These peptides can specifically inhibit or block the binding of MDM2 to
 CC the human p53 protein, in vitro or in vivo. Inhibiting the interaction
 CC between the p53 and MDM2 can induce growth arrest or apoptosis in tumour
 CC cells comprising a wild-type p53 and non-elevated levels of MDM2. The
 CC peptides may be used to identify molecules that bind to MDM2 and to
 CC identify and design inhibitors of MDM2/p53 binding. They may also be
 CC used to purify binding partners especially MDM2, diagnose disease by
 CC measuring levels of MDM2 in blood of cancer and leukaemia patients and
 CC for treatment or prevention of disease involving p53/MDM2 interactions,
 CC especially tumours and viral infections. The peptides can be administered
 CC nasally, rectally, orally or by injection. By interfering with MDM2/p53
 CC interaction, the peptides can activate p53 function and accumulation in
 CC normal cells. The peptides which mimic the MDM2 binding site in p53, have
 CC a significantly greater blocking activity compared with wild-type p53.

XX Sequence 15 AA:

AAW37225 Length: 15 May 30, 2002 09:25 Type: P Check: 8833 ..

1 PAFSFRWSDL SAGAH

11AA_SEQUENCE 1.0
 ID AAW37226 standard; peptide: 10 AA.

AC AAW37226;

```

DE 20-JUL-1998 (first entry)
XX
XX MDM2 binding peptide phage consensus sequence.
DE
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX MDM2; diagnosis; binding; viral infection; phage insert.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX
XX Misc-difference 2 /note= "any natural amino acid"
XX
XX Misc-difference 4 /note= "any natural amino acid"
XX
XX Misc-difference 8 /note= "any natural amino acid"
XX
XX Misc-difference 9 /note= "any natural amino acid"
XX
XX Misc-difference 9 /note= "any natural amino acid"
XX
XX WO9801467-A2.
XX
XX 15-JAN-1998.
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX 07-APR-1997; 97GB-0007041.
XX
XX 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX
XX (NOVS ) NOVARTIS AG.
XX
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
XX Hockeppel H, Lane D, Plicksley S;
XX
XX WPI; 1998-100996/09.
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
XX p53 - useful in, e.g. diagnosis and treatment of cancer and viral
XX infections and identifying binding agents
XX
XX Claim 5; Page 41; 45pp; English.
XX
XX This is a phage consensus sequence of the MDM2 binding peptide identified
XX by phage display. This shows similarity to the known MDM2 binding motif
XX on p53 (TRESLM). The MDM2 binding peptides and their derivatives are
XX capable of binding to the human oncogenic protein MDM2. These peptides
XX can specifically inhibit or block the binding of MDM2 to the human p53
XX protein, in vitro or in vivo. Inhibiting the interaction between the p53
XX and MDM2 can induce growth arrest or apoptosis in tumour cells comprising
XX a wild-type p53 and non-elevated levels of MDM2. The peptides may be used
XX to identify molecules that bind to MDM2 and to identify and design
XX inhibitors of MDM2/p53 binding. They may also be used to purify binding
XX partners especially MDM2, diagnose disease by measuring levels of MDM2 in
XX blood of cancer and leukaemia patients and for treatment or prevention of
XX disease involving p53/MDM2 interactions, especially tumours and viral
XX infections. The peptides can be administered nasally, rectally, orally or
XX by injection. By interfering with MDM2/p53 interaction, the peptides can
XX activate p53 function and accumulation in normal cells. The peptides
XX which mimic the MDM2 binding site in p53, have a significantly greater
XX blocking activity compared with wild-type p53.
XX
XX Sequence 10 AA:
XX
XX AAM37226 Length: 10 May 30, 2002 09:25 Type: P Check: 4557 ..
XX
XX 1 PXXFDYWXXL
XX
XX !:AA_SEQUENCE 1.0
XX
XX ID AAM37195 standard; peptide; 12 AA.
XX
XX AC AAM37195;
XX
XX AC 20-JUL-1998 (first entry)
XX
XX DT

```

```

XX
XX Human oncogenic protein MDM2 binding C-amidated peptide derivative 1.
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX MDM2; diagnosis; binding; viral infection.
XX
XX Synthetic.
XX
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX
XX Modified-site 12 /note= "C-terminal amide"
XX
XX WO9801467-A2.
XX
XX 15-JAN-1998.
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX 07-APR-1997; 97GB-0007041.
XX
XX 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX
XX (NOVS ) NOVARTIS AG.
XX
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
XX Hockeppel H, Lane D, Plicksley S;
XX
XX WPI; 1998-100996/09.
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
XX p53 - useful in, e.g. diagnosis and treatment of cancer and viral
XX infections and identifying binding agents
XX
XX Example 1; Page 20; 45pp; English.
XX
XX This is a C-amidated peptide derivative capable of binding to a human
XX oncogenic protein MDM2. The MDM2 binding peptides can specifically
XX inhibit or block the binding of MDM2 to the human p53 protein, in vitro
XX or in vivo. Inhibiting the interaction between the p53 and MDM2 can
XX induce growth arrest or apoptosis in tumour cells comprising a wild-type
XX p53 and non-elevated levels of MDM2. The peptides may be used to identify
XX molecules that bind to MDM2 and to identify and design inhibitors of
XX MDM2/p53 binding. They may also be used to purify binding partners
XX especially MDM2, diagnose disease by measuring levels of MDM2 in blood
XX of cancer and leukaemia patients and for treatment or prevention of
XX disease involving p53/MDM2 interactions, especially tumours and viral
XX infections. The peptides can be administered nasally, rectally, orally
XX or by injection. By interfering with MDM2/p53 interaction, the peptides
XX can activate p53 function and accumulation in normal cells. The peptides
XX which mimic the MDM2 binding site in p53, have a significantly greater
XX blocking activity compared with wild-type p53.
XX
XX Sequence 12 AA:
XX
XX AAM37195 Length: 12 May 30, 2002 09:25 Type: P Check: 5993 ..
XX
XX 1 TGPATHTYMA TF
XX
XX !:AA_SEQUENCE 1.0
XX
XX ID AAM37196 standard; peptide; 12 AA.
XX
XX AC AAM37196;
XX
XX AC 20-JUL-1998 (first entry)
XX
XX DE Human oncogenic protein MDM2 binding C-amidated peptide derivative 2.
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX MDM2; diagnosis; binding; viral infection.
XX
XX Synthetic.
XX
XX OS Homo sapiens.
XX

```

```

XX Key Location/Qualifiers
FH Modified-site 12
FT /note- "C-terminal amide"
XX
XX WO9801467-A2.
XX
XX 15-JAN-1998.
XX
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX
XX 07-APR-1997; 97GB-0007041.
XX
XX 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX (NOVS ) NOVARTIS AG.
XX
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
XX Hochkeppel H, Lane D, Picklesley S;
XX
XX WPI; 1998-100996/09.
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
XX p53 - useful in, e.g. diagnosis and treatment of cancer and viral
XX infections and identifying binding agents
XX
XX Example 1; Page 20; 45pp; English.
XX
XX This is a C-amidated peptide derivative capable of binding to a human
XX oncogenic protein MDM2. The MDM2 binding peptides can specifically
XX inhibit or block the binding of MDM2 to the human p53 protein, in vitro
XX or in vivo. Inhibiting the interaction between the p53 and MDM2 can
XX induce growth arrest or apoptosis in tumour cells comprising a wild-type
XX p53 and non-elevated levels of MDM2. The peptides may be used to identify
XX molecules that bind to MDM2 and to identify and design inhibitors of
XX MDM2/p53 binding. They may also be used to purify binding partners
XX especially MDM2, diagnose disease by measuring levels of MDM2 in blood
XX of cancer and leukemia patients and for treatment or prevention of
XX disease involving p53/MDM2 interactions, especially tumours and viral
XX infections. The peptides can be administered nasally, rectally, orally
XX or by injection. By interfering with MDM2/p53 interaction, the peptides
XX can activate p53 function and accumulation in normal cells. The peptides
XX which mimic the MDM2 binding site in p53, have a significantly greater
XX blocking activity compared with wild-type p53.
XX
XX Sequence 12 AA;
XX
XX AAW37196 Length: 12 May 30, 2002 09:25 Type: P Check: 5978 ..
XX
XX 1 MPREMDYMEG LN
XX
XX 11AA_SEQUENCE 1.0
XX ID AAW37197 standard; peptide; 14 AA.
XX
XX AAW37197;
XX
XX 20-JUL-1998 (first entry)
XX
XX Human oncogenic protein MDM2 binding Cys (Acrid) peptide derivative 1.
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX tumour; diagnosis; binding; viral infection.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH Modified-site 1
XX FT /note- "N-terminal acetyl; Sulphydryl side-chain
XX FT linked to 6-acryloyl-2-(dimethylamino)
XX FT naphthalene"
XX
XX Modified-site 14
XX /note- "C-terminal amide"
XX

```

```

XX
XX WO9801467-A2.
XX
XX 15-JAN-1998.
XX
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX
XX 07-APR-1997; 97GB-0007041.
XX
XX 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX (NOVS ) NOVARTIS AG.
XX
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
XX Hochkeppel H, Lane D, Picklesley S;
XX
XX WPI; 1998-100996/09.
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
XX p53 - useful in, e.g. diagnosis and treatment of cancer and viral
XX infections and identifying binding agents
XX
XX Example 2; Page 20; 45pp; English.
XX
XX This is a Cys (Acrid) peptide derivative capable of binding to a human
XX oncogenic protein MDM2. The MDM2 binding peptides can specifically
XX inhibit or block the binding of MDM2 to the human p53 protein, in vitro
XX or in vivo. Inhibiting the interaction between the p53 and MDM2 can
XX induce growth arrest or apoptosis in tumour cells comprising a wild-type
XX p53 and non-elevated levels of MDM2. The peptides may be used to identify
XX molecules that bind to MDM2 and to identify and design inhibitors of
XX MDM2/p53 binding. They may also be used to purify binding partners
XX especially MDM2, diagnose disease by measuring levels of MDM2 in blood
XX of cancer and leukemia patients and for treatment or prevention of
XX disease involving p53/MDM2 interactions, especially tumours and viral
XX infections. The peptides can be administered nasally, rectally, orally
XX or by injection. By interfering with MDM2/p53 interaction, the peptides
XX can activate p53 function and accumulation in normal cells. The peptides
XX which mimic the MDM2 binding site in p53, have a significantly greater
XX blocking activity compared with wild-type p53.
XX
XX Sequence 14 AA;
XX
XX AAW37197 Length: 14 May 30, 2002 09:25 Type: P Check: 8253 ..
XX
XX 1 CGPTEFSDYW KLP
XX
XX 11AA_SEQUENCE 1.0
XX ID AAW37203 standard; peptide; 16 AA.
XX
XX AAW37203;
XX
XX 20-JUL-1998 (first entry)
XX
XX Human oncogenic protein MDM2 binding biotinylated peptide derivative 3.
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX tumour; diagnosis; binding; viral infection; biotinylation.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH Modified-site 1
XX FT /note- "biotinylated"
XX FT Modified-site 16
XX FT /note- "C-terminal amide"
XX
XX WO9801467-A2.
XX
XX 15-JAN-1998.
XX
XX 04-JUL-1997; 97WO-EP03549.
XX

```

```
XX 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX (NOVS ) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
XX WPI: 1998-100996/09.
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
XX Example 3; Page 21; 45pp; English.
XX
XX This is a biotinylated peptide derivative capable of binding to a human
XX oncogenic protein MDM2. The MDM2 binding peptides can specifically
XX inhibit or block the binding of MDM2 to the human p53 protein, in vitro
XX or in vivo. Inhibiting the interaction between the p53 and MDM2 can
XX induce growth arrest or apoptosis in tumour cells comprising a wild-type
XX p53 and non-elevated levels of MDM2. The peptides may be used to identify
XX molecules that bind to MDM2 and to identify and design inhibitors of
XX MDM2/p53 binding. They may also be used to purify binding partners
XX especially MDM2, diagnose disease by measuring levels of MDM2 in blood
XX of cancer and leukaemia patients and for treatment or prevention of
XX disease involving p53/MDM2 interactions, especially tumours and viral
XX infections. The peptides can be administered nasally, rectally, orally
XX or by injection. By interfering with MDM2/p53 interaction, the peptides
XX can activate p53 function and accumulation in normal cells. The peptides
XX which mimic the MDM2 binding site in p53, have a significantly greater
XX blocking activity compared with wild-type p53.
XX
XX Sequence 16 AA:
AAW37203 Length: 16 May 30, 2002 09:25 Type: P Check: 634 ..
1 SSGSQETFSQ YWKLIP
11AA_SEQUENCE 1.0
ID AAW37204 standard; peptide; 28 AA.
XX
AC AAW37204;
XX
XX 20-JUL-1998 (first entry)
XX
XX Human oncogenic protein MDM2 binding biotinylated peptide derivative 4.
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX tumour; diagnosis; binding; viral infection; biotinylation.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 1 /note= "biotinylated"
XX Modified-site 28 /note= "C-terminal amide"
XX
XX W09801467-A2.
XX
XX 15-JAN-1998.
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX 07-APR-1997; 97GB-0007041.
XX 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX (NOVS ) NOVARTIS AG.
PA
```

```
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picksley S;
XX
XX WPI: 1998-100996/09.
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
XX Example 3; Page 21; 45pp; English.
XX
XX This is a biotinylated peptide derivative capable of binding to a human
XX oncogenic protein MDM2. The MDM2 binding peptides can specifically
XX inhibit or block the binding of MDM2 to the human p53 protein, in vitro
XX or in vivo. Inhibiting the interaction between the p53 and MDM2 can
XX induce growth arrest or apoptosis in tumour cells comprising a wild-type
XX p53 and non-elevated levels of MDM2. The peptides may be used to identify
XX molecules that bind to MDM2 and to identify and design inhibitors of
XX MDM2/p53 binding. They may also be used to purify binding partners
XX especially MDM2, diagnose disease by measuring levels of MDM2 in blood
XX of cancer and leukaemia patients and for treatment or prevention of
XX disease involving p53/MDM2 interactions, especially tumours and viral
XX infections. The peptides can be administered nasally, rectally, orally
XX or by injection. By interfering with MDM2/p53 interaction, the peptides
XX can activate p53 function and accumulation in normal cells. The peptides
XX which mimic the MDM2 binding site in p53, have a significantly greater
XX blocking activity compared with wild-type p53.
XX
XX Sequence 28 AA:
AAW37204 Length: 28 May 30, 2002 09:25 Type: P Check: 1571 ..
1 SMPRFMDYWE GINROIKIWF QNRMRKWK
11AA_SEQUENCE 1.0
ID AAW37205 standard; peptide; 31 AA.
XX
AC AAW37205;
XX
XX 20-JUL-1998 (first entry)
XX
XX Human oncogenic protein MDM2 binding biotinylated peptide derivative 5.
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX tumour; diagnosis; binding; viral infection; biotinylation.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 1 /note= "N-terminal acetyl"
XX Modified-site 17 /label= bala
XX Modified-site 30 /note= "beta-Alanine"
XX Modified-site 30 /label= bala
XX Modified-site 31 /note= "beta-Alanine"
XX Modified-site 31 /note= "biotinylated"
XX
XX W09801467-A2.
XX
XX 15-JAN-1998.
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX 07-APR-1997; 97GB-0007041.
XX 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX (NOVS ) NOVARTIS AG.
PA
```

XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picklesley S;
XX
DR WPI: 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 3; Page 21; 45pp; English.
XX
CC This is a biotinylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein. In vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 31 AA;
AAW37205 Length: 31 May 30, 2002 09:25 Type: P Check: 8233 ..
1 AAWALPAVL IALLAPXMPR FMDYEGLINX K
11AA_SEQUENCE 1.0
ID AAW37182 standard; peptide; 12 AA.
XX
AC AAW37182;
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding N-acetylated peptide derivative 1.
XX
KM MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KM tumour; diagnosis; binding; viral infection.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 12 /note= "C-terminal amide"
FT
FT
XX WO9801467-A2.
XX
PN 15-JAN-1998.
XX
PD
XX
PF 04-JUL-1997; 97WO-EP03549.
XX
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picklesley S;
XX
XX WPI: 1998-100996/09.
XX

PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 1; Page 18; 45pp; English.
XX
CC This is a N-acetylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein. In vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 12 AA;
AAW37182 Length: 12 May 30, 2002 09:25 Type: P Check: 5993 ..
1 TGPAFTHYMA TF
11AA_SEQUENCE 1.0
ID AAW37183 standard; peptide; 12 AA.
XX
AC AAW37183;
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding N-acetylated peptide derivative 2.
XX
KM MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KM tumour; diagnosis; binding; viral infection.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 12 /note= "C-terminal amide"
FT
FT
XX WO9801467-A2.
XX
PN 15-JAN-1998.
XX
PD
XX
PF 04-JUL-1997; 97WO-EP03549.
XX
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picklesley S;
XX
XX WPI: 1998-100996/09.
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 1; Page 19; 45pp; English.
XX

CC This is a N-acylated peptide derivative capable of binding to a human
 CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
 CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
 CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
 CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
 CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
 CC molecules that bind to MDM2 and to identify and design inhibitors of
 CC MDM2/p53 binding. They may also be used to purify binding partners
 CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
 CC of cancer and leukaemia patients and for treatment or prevention of
 CC disease involving p53/MDM2 interactions, especially tumours and viral
 CC infections. The peptides can be administered nasally, rectally, orally
 CC or by injection. By interfering with MDM2/p53 interaction, the peptides
 CC can activate p53 function and accumulation in normal cells. The peptides
 CC which mimic the MDM2 binding site in p53, have a significantly greater
 CC blocking activity compared with wild-type p53.

SQ Sequence 12 AA:

AAW37183 Length: 12 May 30, 2002 09:25 Type: P Check: 5978 ..

1 MPRFMDYMEG LN

IIAA_SEQUENCE 1.0
 ID AAW37184 standard; peptide; 12 AA.
 AC AAW37184;
 XX
 XX
 DF 20-JUL-1998 (first entry)
 DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 3.
 XX
 XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
 KM tumour; diagnosis; binding; viral infection.
 OS Synthetic.
 OS Homo sapiens.
 FT Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 12
 FT /note= "C-terminal amide"
 XX
 PN W09801467-A2.
 PD 15-JAN-1998.
 XX
 XX 04-JUL-1997; 97WO-EP03549.
 PF
 PR 07-APR-1997; 97GB-0007041.
 PR 05-JUL-1996; 96GB-0014197.
 XX
 PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
 PA (NOVS) NOVARTIS AG.
 XX
 XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
 PI Hochkeppel H, Lane D, Pickersley S;
 XX
 XX WPI: 1998-100996/09.
 DR
 XX
 PT Compounds binding to MDM2 protein and inhibit its interaction with
 PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
 PT infections and identifying binding agents
 XX
 XX Example 1; Page 19; 45pp; English.
 XX
 CC This is a N-acylated peptide derivative capable of binding to a human
 CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
 CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
 CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
 CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
 CC p53 and non-elevated levels of MDM2. The peptides may be used to identify

CC molecules that bind to MDM2 and to identify and design inhibitors of
 CC MDM2/p53 binding. They may also be used to purify binding partners
 CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
 CC of cancer and leukaemia patients and for treatment or prevention of
 CC disease involving p53/MDM2 interactions, especially tumours and viral
 CC infections. The peptides can be administered nasally, rectally, orally
 CC or by injection. By interfering with MDM2/p53 interaction, the peptides
 CC can activate p53 function and accumulation in normal cells. The peptides
 CC which mimic the MDM2 binding site in p53, have a significantly greater
 CC blocking activity compared with wild-type p53.

SQ Sequence 12 AA:

AAW37184 Length: 12 May 30, 2002 09:25 Type: P Check: 6146 ..

1 QPFSDYWKL LP

IIAA_SEQUENCE 1.0
 ID AAW37190 standard; peptide; 12 AA.
 AC AAW37190;
 XX
 XX
 DF 20-JUL-1998 (first entry)
 DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 9.
 XX
 XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
 KM tumour; diagnosis; binding; viral infection.
 OS Synthetic.
 OS Homo sapiens.
 FT Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 12
 FT /note= "C-terminal amide"
 XX
 PN W09801467-A2.
 PD 15-JAN-1998.
 XX
 XX 04-JUL-1997; 97WO-EP03549.
 PF
 PR 07-APR-1997; 97GB-0007041.
 PR 05-JUL-1996; 96GB-0014197.
 XX
 PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
 PA (NOVS) NOVARTIS AG.
 XX
 XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
 PI Hochkeppel H, Lane D, Pickersley S;
 XX
 XX WPI: 1998-100996/09.
 DR
 XX
 PT Compounds binding to MDM2 protein and inhibit its interaction with
 PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
 PT infections and identifying binding agents
 XX
 XX Example 1; Page 19; 45pp; English.
 XX
 CC This is a N-acylated peptide derivative capable of binding to a human
 CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
 CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
 CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
 CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
 CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
 CC molecules that bind to MDM2 and to identify and design inhibitors of
 CC MDM2/p53 binding. They may also be used to purify binding partners
 CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
 CC of cancer and leukaemia patients and for treatment or prevention of
 CC disease involving p53/MDM2 interactions, especially tumours and viral
 CC infections. The peptides can be administered nasally, rectally, orally

CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.

SQ Sequence 12 AA:

AAW37190 Length: 12 May 30, 2002 09:25 Type: P Check: 6124 ..

1 QETFSDWKL LP

11AA_SEQUENCE 1.0
ID AAW37191 standard; peptide; 12 AA.
XX
AC AAW37191:
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 10.
XX
KM MDM2: oncogenic protein; p53: human; inhibition; interaction; cancer;
KM tumour; diagnosis; binding; viral infection.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 12 /note= "C-terminal amide"
XX
XX MO9801467-A2.
XX
PD 15-JAN-1998.
XX
PF 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Plicksley S;
XX
DR WPI; 1998-100996/09.
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
XX Example 1; Page 19; 45pp; English.
XX
CC This is a N-acylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 12 AA:

AAW37191 Length: 12 May 30, 2002 09:25 Type: P Check: 6151 ..

1 VQNFIDYWTQ QF

11AA_SEQUENCE 1.0
ID AAW37192 standard; peptide; 15 AA.
XX
AC AAW37192:
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 11.
XX
KM MDM2: oncogenic protein; p53: human; inhibition; interaction; cancer;
KM tumour; diagnosis; binding; viral infection.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 15 /note= "C-terminal amide"
XX
XX MO9801467-A2.
XX
PD 15-JAN-1998.
XX
PF 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Plicksley S;
XX
DR WPI; 1998-100996/09.
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
XX Example 1; Page 19; 45pp; English.
XX
CC This is a N-acylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 15 AA:
XX
AAW37192 Length: 15 May 30, 2002 09:25 Type: P Check: 9093 ..
XX
XX 1 IDRAPTFRDH WFALY
XX
11AA_SEQUENCE 1.0

ID AAW37193 standard; peptide; 15 AA.
 XX
 AC AAW37193;
 XX
 DT 20-JUL-1998 (first entry)
 XX
 DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 12.
 XX
 KM MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
 KW tumour; diagnosis; binding; viral infection.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site /note= "N-terminal acetyl"
 FT Modified-site 15
 FT Modified-site /note= "C-terminal amide"
 XX
 PN WO9801467-A2.
 PD 15-JAN-1998.
 XX
 PF 04-JUL-1997; 97WO-EP03549.
 XX
 PR 07-APR-1997; 97GB-0007041.
 PR 05-JUL-1996; 96GB-0014197.
 XX
 PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
 PA (NOVS) NOVARTIS AG.
 XX
 PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
 PI Hochkeppel H, Lane D, Picksley S;
 XX
 DR WP1; 1998-100996/09.
 XX
 PS Compounds binding to MDM2 protein and inhibit its interaction with
 PT p53 useful in, e.g. diagnosis and treatment of cancer and viral
 PT infections and identifying binding agents
 XX
 PS Example 1; Page 19; 45pp; English.
 XX
 CC This is a N-acylated peptide derivative capable of binding to a human
 CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
 CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
 CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
 CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
 CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
 CC molecules that bind to MDM2 and to identify and design inhibitors of
 CC MDM2/p53 binding. They may also be used to purify binding partners
 CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
 CC of cancer and leukaemia patients and for treatment or prevention of
 CC disease involving p53/MDM2 interactions, especially tumours and viral
 CC infections. The peptides can be administered nasally, rectally, orally
 CC or by injection. By interfering with MDM2/p53 interaction, the peptides
 CC can activate p53 function and accumulation in normal cells. The peptides
 CC which mimic the MDM2 binding site in p53, have a significantly greater
 CC blocking activity compared with wild-type p53.
 XX
 SO Sequence 15 AA;
 AAW37193 Length: 15 May 30, 2002 09:25 Type: P Check: 9428 ..
 1 PRPALVFADY WENTLY
 !IAA_SEQUENCE 1.0
 ID AAW37194 standard; peptide; 15 AA.
 XX
 AC AAW37194;
 XX
 DT 20-JUL-1998 (first entry)
 XX

DE Human oncogenic protein MDM2 binding N-acylated peptide derivative 13.
 XX
 KM MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
 KW tumour; diagnosis; binding; viral infection.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site /note= "N-terminal acetyl"
 FT Modified-site 15
 FT Modified-site /note= "C-terminal amide"
 XX
 PN WO9801467-A2.
 PD 15-JAN-1998.
 XX
 PF 04-JUL-1997; 97WO-EP03549.
 XX
 PR 07-APR-1997; 97GB-0007041.
 PR 05-JUL-1996; 96GB-0014197.
 XX
 PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
 PA (NOVS) NOVARTIS AG.
 XX
 PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
 PI Hochkeppel H, Lane D, Pickesley S;
 XX
 DR WP1; 1998-100996/09.
 XX
 PS Compounds binding to MDM2 protein and inhibit its interaction with
 PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
 PT infections and identifying binding agents
 XX
 PS Example 1; Page 19; 45pp; English.
 XX
 CC This is a N-acylated peptide derivative capable of binding to a human
 CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
 CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
 CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
 CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
 CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
 CC molecules that bind to MDM2 and to identify and design inhibitors of
 CC MDM2/p53 binding. They may also be used to purify binding partners
 CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
 CC of cancer and leukaemia patients and for treatment or prevention of
 CC disease involving p53/MDM2 interactions, especially tumours and viral
 CC infections. The peptides can be administered nasally, rectally, orally
 CC or by injection. By interfering with MDM2/p53 interaction, the peptides
 CC can activate p53 function and accumulation in normal cells. The peptides
 CC which mimic the MDM2 binding site in p53, have a significantly greater
 CC blocking activity compared with wild-type p53.
 XX
 SO Sequence 15 AA;
 AAW37194 Length: 15 May 30, 2002 09:25 Type: P Check: 8633 ..
 1 PAFSRFMSDL SAGAH
 !IAA_SEQUENCE 1.0
 ID AAW37170 standard; peptide; 12 AA.
 XX
 AC AAW37170;
 XX
 DT 20-JUL-1998 (first entry)
 XX
 DE Human oncogenic protein MDM2 binding peptide 1.
 XX
 KM MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
 KW tumour; diagnosis; binding; viral infection.
 XX
 OS Synthetic.

OS Homo sapiens.
XX
XX WO9801467-A2.
XX
XX 15-JAN-1998.
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picklesley S;
XX
XX WPI; 1998-100996/09.
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
XX Claim 5; Page 41; 45pp; English.
XX
XX This peptide is capable of binding to an oncogenic protein MDM2
CC (especially human DM2). The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
XX Sequence 12 AA;
SQ
AAW37170 Length: 12 May 30, 2002 09:25 Type: P Check: 5978 ..
1 MPRMDYWG LN
11AA_SEQUENCE 1.0
ID AAW37171 standard; peptide; 12 AA.
XX
XX AAW37171;
XX
XX 20-JUL-1998 (first entry)
XX
XX Human oncogenic protein MDM2 binding peptide 2.
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX tumour; diagnosis; binding; viral infection.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX WO9801467-A2.
XX
XX 15-JAN-1998.
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA

PA (NOVS) NOVARTIS AG.
XX
XX Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picklesley S;
XX
XX WPI; 1998-100996/09.
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
XX Claim 5; Page 41; 45pp; English.
XX
XX This peptide is capable of binding to an oncogenic protein MDM2
CC (especially human DM2). The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukaemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
XX Sequence 12 AA;
SQ
AAW37171 Length: 12 May 30, 2002 09:25 Type: P Check: 6146 ..
1 QPTSDYWKL LP
11AA_SEQUENCE 1.0
ID AAW37172 standard; peptide; 15 AA.
XX
XX AAW37172;
XX
XX 20-JUL-1998 (first entry)
XX
XX Human oncogenic protein MDM2 binding peptide 3.
XX
XX MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
XX tumour; diagnosis; binding; viral infection.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX WO9801467-A2.
XX
XX 15-JAN-1998.
XX
XX 04-JUL-1997; 97WO-EP03549.
XX
XX 07-APR-1997; 97GB-0007041.
PR 05-JUL-1996; 96GB-0014197.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
PA (NOVS) NOVARTIS AG.
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Picklesley S;
XX
XX WPI; 1998-100996/09.
XX
XX Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
XX Disclosure; Page 4; 45pp; English.
PS

XX This peptide is capable of binding to an oncogenic protein MDM2
 CC (especially human DM2). The MDM2 binding peptides can specifically
 CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
 CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
 CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
 CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
 CC molecules that bind to MDM2 and to identify and design inhibitors of
 CC MDM2/p53 binding. They may also be used to purify binding partners
 CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
 CC of cancer and leukemia patients and for treatment or prevention of
 CC disease involving p53/MDM2 interactions, especially tumours and viral
 CC infections. The peptides can be administered nasally, rectally, orally
 CC or by injection. By interfering with MDM2/p53 interaction, the peptides
 CC can activate p53 function and accumulation in normal cells. The peptides
 CC which mimic the MDM2 binding site in p53, have a significantly greater
 CC blocking activity compared with wild-type p53.

XX Sequence 15 AA;

AAW37172 Length: 15 May 30, 2002 09:25 Type: P Check: 9428 ..

1 PRPALVFADY WETLY

!!AA_SEQUENCE 1.0
 ID AAW46748 standard; Protein; 534 AA.

XX AAW46748;

XX 06-JUL-1998 (first entry)

XX Arabidopsis thaliana phosphate transporter 2 protein.

XX Phosphate transporter protein; AtPT1; AtPT2; LePT1; LePT2;
 KW phosphate-deficiency.

XX Arabidopsis thaliana.

XX Key Location/Qualifiers

FT Misc-difference 290 /note= "encoded by TTC"

FT Misc-difference 437 /note= "encoded by TTC"

FT Misc-difference 437 /note= "encoded by TTC"

XX WO9804701-A1.

XX 05-FEB-1998.

XX 29-JUL-1997; 97WO-US13458.

XX 29-JUL-1996; 96US-0022391.

XX (PURD) PURDUE RES FOUND.

XX Muchhal US, Raghothama KG;

XX MPI: 1998-130690/12.

XX N-PSDB; AAV16291.

XX Plant phosphate transporter proteins - used to improve plants
 PT ability to grow in phosphate-deficient soil

XX Claim 2; Page 38-40; 66pp; English.

XX This sequence is the Arabidopsis thaliana phosphate transporter
 CC protein 2 (AtPT2). AtPT1 (see AAW46747) AtPT2 (see AAW46748), LePT1 (see
 CC AAW46749), and LePT2 (see AAW46750) are the proteins of the invention.
 CC AtPT1 and AtPT2 are from Arabidopsis thaliana, while LePT1 and LePT2 are
 CC the tomato equivalents of the Arabidopsis proteins. The DNA sequences,
 CC and constructs and vectors containing them can be used for improving a
 CC plant's ability to grow in phosphate-deficient soil.

XX Sequence 534 AA;

AAW46748 Length: 534 May 30, 2002 09:25 Type: P Check: 562 ..
 1 MARQLOVLN ALDVAKTQWY HFTALIIAGM GFTDADLFC CISLVTKLLG
 51 RIYHVGAQG KPELTLPNNA AAVNGVAFG TLAQLEFFGW LGKLRKKV
 101 YGMLVMVVL CSTASGISFG HEPRKVMATL CFFRFWISFG IGGDYPISAT
 151 IMSEYANKKT RCAFVSAYFA MCGFGIMAGG IFALIIISAF EAKFPSPAVA
 201 DDALGSTRPQ ADLVWRILIM AGAIRAAMY YSRKMEETA RYALYAKDA
 251 KQASDMSKV LOVEIEPEQO KLEISKEKS KAFGLFSKEE MSRHGLHLG
 301 TTSTWPLLDI AFYSQNLFOK DIFSAIGWIP PAQSMNATQE VFKIRAOQL
 351 IALCSTVPGY WFTVAFIDVI GRFALQMGF FFMVVEFAL AIYNNWTHK
 401 ENRIGEVIMY SLTEFFANFG PNATTFVVA EIPPARERST CHGISASGK
 451 LGAMVGAFGF LYLAQNPDKD KTDAGYPPGI GVRNSLIYIG VVNFGLILFT
 501 FIVPEKSKGS LDEMSEGENE NENSNDSTRT VPIV

!!AA_SEQUENCE 1.0
 ID AAW46749 standard; Protein; 538 AA.

XX AAW46749;

XX 06-JUL-1998 (first entry)

XX Tomato phosphate transporter 1 protein.

XX Phosphate transporter protein; AtPT1; AtPT2; LePT1; LePT2; tomato;
 KW phosphate-deficiency.

XX Lycopersicon esculentum.

XX WO9804701-A1.

XX 05-FEB-1998.

XX 29-JUL-1997; 97WO-US13458.

XX 29-JUL-1996; 96US-0022391.

XX (PURD) PURDUE RES FOUND.

XX Muchhal US, Raghothama KG;

XX MPI: 1998-130690/12.

XX N-PSDB; AAV16292.

XX Plant phosphate transporter proteins - used to improve plants
 PT ability to grow in phosphate-deficient soil

XX Claim 2; Page 41-44; 66pp; English.

XX This sequence is the Lycopersicon esculentum (tomato) phosphate
 CC transporter protein 1 (LePT1). AtPT1 (see AAW46747) AtPT2 (see
 CC AAW46748), LePT1 (see AAW46749), and LePT2 (see AAW46750) are the
 CC proteins of the invention. AtPT1 and AtPT2 are from Arabidopsis thaliana,
 CC while LePT1 and LePT2 are the tomato equivalents of the Arabidopsis
 CC proteins. The DNA sequences, and constructs and vectors containing them
 CC can be used for improving a plant's ability to grow in
 CC phosphate-deficient soil.

XX Sequence 538 AA;

AAW46749 Length: 538 May 30, 2002 09:25 Type: P Check: 4793 ..

1 MANDLOVLNA LDVAKTQLYH FTAIYIAGMG FFDVADLDFC ISWYTKLLGR

51 LYHHHDAK PGSLPPNVSA AVNGVAFQST LAGOLFEGWL GDRKGRKKV
101 GMTIMIVIC SIASGLSFCH TRKGWTTIC FFRFWLGRGI GGGYPLSATI
151 MSEYANKTR GAFIAAVFAM QGFGIAGM VAIIVSAFK GAPPAPAYEV
201 DAIGSTVPOA DFVWRIILME GAIPAGLTYY WRMKPEFAR YTAIVANLKL
251 QNANDMSKVL QVEIEAPEK VTAISEAKA NDFGLTFKEF LRRHGLLLG
301 TASTWFLDI AFYSQNLFOK DIFSAIGWIP PAQTMNALEE VKIARAQTL
351 IALCSTVPGY WETVAFIDKI GRPAIOLMGF FFMVTFMEAL AIDYHMTLK
401 DHRIGFVVM STFFFFANEG PNATTFVYPA EIPPARLST CHGISAAAGK
451 AGANVGAFFG LYAQPDPPT KTDAGYPPGI GVRNSLIYLG CVNFGMLFT
501 FLVPSNGKS LEDLSRENEG EETVAEIRA TSGRTVPV
11AA_SEQUENCE 1.0
ID AAM46750 standard; Protein: 528 AA.
AC AAM46750;
XX
XX
DT 06-JUL-1998 (first entry)
DE Tomato phosphate transporter 1 protein.
XX
XX Phosphate transporter protein; AtPT1; AtPT2; LepT1; LepT2; tomato;
KW phosphate-deficiency.
XX
XX Lycopersicon esculentum.
OS
FH Key Location/Qualifiers
FT Misc-difference 364
FT /note= "encoded by GCG"
XX
XX MO9804701-A1.
XX
XX PD 05-FEB-1998.
XX
XX PF 29-JUL-1997; 97MO-US13458.
XX
XX PR 29-JUL-1996; 96US-0022391.
XX
XX PA (PURD) PURDUE RES FOUND.
XX
XX PI Muchhal US, Raghochama KG;
XX
XX DR WPI: 1998-130690/12.
XX
XX DR N-PSDB; AAV16293.
XX
XX PT Plant phosphate transporter proteins - used to improve plants
XX
XX PT ability to grow in phosphate-deficient soil
XX
XX PS Claim 2; Page 41-44; 66pp; English.
XX
XX This sequence is the Lycopersicon esculentum (tomato) phosphate
XX transporter protein 1 (LepT1). AtPT1 (see AAM46747) AtPT2 (see
XX AAM46748), LepT1 (see AAM46749), and LepT2 (see AAM46750) are the
XX proteins of the invention. AtPT1 and AtPT2 are from Arabidopsis thaliana,
XX while LepT1 and LepT2 are the tomato equivalents of the Arabidopsis
XX proteins. The DNA sequences, and constructs and vectors containing them
XX can be used for improving a plant's ability to grow in
XX phosphate-deficient soil.
XX
XX Sequence 528 AA;
AAM46750 Length: 528 May 30, 2002 09:25 Type: P Check: 3154 ..
1 MAVGDNDNNN LQVINALDLA KTQLYHTTAI VIAGMGFTTD AYDLFSISLV

51 TKLGRLYXT KPDLLKPGL PPAVSATYG VALVGTLAGQ LFFGWLGDKM
101 GRKKYVGMTL VLVAVCSVAS GLSFGSTPKG VMTILCFRRF WLGCIGIGDY
151 PLSATIMSEY ANKTRGAFI AAVFAMQCGF ILFSGIVALI TAAQFDHAYR
201 SPTEENNAL STVPQSDYIW RIILMGSLP AALTYYMRK MPETARYTAL
251 VAKDAKRAQ DMGKYLQVEI ESEEAKEIOI SRNEINQGL FSWEFVRRHG
301 LHFGTCSTW FLDDIAFYSQ NLFQKDVFA VGMIPKAPTM NAVQEVYKIA
351 RAQFLIALCS TVPCYMETVA FIDIIGRFPI QLMGFEMTV FMAIAPYH
401 HMTLEANNIG FIWYSLTFF FANEGNALT FVYPAEIRPA RLKSTGIGIS
451 AAGKAGAIY GAVGFLYAAQ SKDPNKTGAG YPAGIGIKNS LIVIGCINAL
501 GMLCTFCVPE PKGSLREAS QETTGEA
11AA_SEQUENCE 1.0
ID AAM37145 standard; Protein: 130 AA.
AC AAM37145;
XX
XX
DT 18-JUN-1998 (first entry)
DE Human O289_1 protein.
XX
XX Human; O289_1; secreted protein; immunomodulator; antitumor;
KW tissue growth; haemostatic; thrombolytic; cancer.
XX
XX OS Homo sapiens.
XX
XX PN MO9801552-A2.
XX
XX PD 15-JAN-1998.
XX
XX PF 07-JUL-1997; 97MO-US11842.
XX
XX PR 27-SEP-1996; 96US-0721488.
XX
XX PR 09-JUL-1996; 96US-0677231.
XX
XX PA (GEMV) GENERICS INST INC.
XX
XX PI Bowman M, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
XX
XX PI Racie LA, Spauldingv, Treacy M;
XX
XX DR WPI: 1998-101050/09.
XX
XX DR N-PSDB; AAV02990.
XX
XX PT New nucleic acid encoding secreted human proteins, and related
XX
XX PT vectors, transformed cells and proteins - are potentially useful
XX
XX PT e.g. as immuno-modulators, antitumor agents, promoters of tissue
XX
XX PT growth, haemostatic and thrombolytic agents etc.
XX
XX PS Claim 17; Page 45; 55pp; English.
XX
XX The present sequence represents the human O289_1 protein, which
XX is a secreted protein. The DNA encoding the secreted protein may be used
XX e.g. to express the recombinant protein; as tissue/molecular weight
XX marker; for chromosome identification; to identify possible genetic
XX disorders; to isolate new related DNA; as sources of primers for
XX polymerase chain reaction; to generate anti-protein or anti-DNA
XX antibodies or in interaction trap assays to identify sequences that
XX encode interacting proteins. The protein can be used to screen compounds
XX for biological activity; to raise antibodies; as tissue markers; for
XX isolation of related receptors and ligands and as nutritional sources.
XX It may also have many biological activities, e.g. cytokine and cell
XX proliferation/differentiation activity; immunosuppressant/immunostimulant
XX activity; regulation of haematopoiesis; inhibition or activation of
XX follicle-stimulating hormone; chemotactic/chemokinetic activity;

CC haemostatic and thrombolytic activity; as receptors or ligands;
 CC antiinflammatory activity; as antimicrobials; modulators of metabolism or
 CC behaviour; as analgesics; enzymes for treating specific deficiency
 CC disorders, and in treatment of psoriasis. Neutralising antibodies
 CC against the protein can be used therapeutically, e.g. to detect or
 CC prevent metastasis of the protein-expressing cancers.

XX Sequence 130 AA;

AAW37145 Length: 130 May 30, 2002 09:25 Type: P Check: 5772 ..

1 NENPGVCHSS ATGDPHCROP EPTSPGPI SM LFSLLPPPR LATHHETSQ

51 LGVQDRDFRV FLSKERDGF YGSLPVGILL AHRVDHTKSC QFGFLSIPTP

101 SPRVRYTALI QALVISHLCH FSEFWIGLPA

IIAA_SEQUENCE 1.0

ID AAY30723 standard; Protein; 182 AA.

XX AAY30723;

DT 18-NOV-1999 (first entry)

XX Amino acid sequence of a human secreted protein.

DE Secreted protein; gene therapy; cancer; tumor; fetal deficiency;
 XX neurodegenerative disorder; developmental abnormality; blood disorder;
 XX immune system disease; autoimmune disease; leukemia; inflammation;
 XX allergy; Alzheimer's disease; cognitive disorder; schizophrenia;
 XX obesity; osteoporosis; arthritis; infection; AIDS; diabetes; asthma;
 XX connective tissue disorder; transplant rejection; sepsis; acne;
 XX psoriasis; cardiovascular disorder; reproductive disorder;
 XX food additive; food preservative; storage capability.

XX Homo sapiens.

XX W09943693-A1.

XX 02-SEP-1999.

XX 24-FEB-1999; 99WO-US03939.

XX 26-FEB-1998; 98US-0076051.

XX 26-FEB-1998; 98US-0076052.

XX 26-FEB-1998; 98US-0076053.

XX 26-FEB-1998; 98US-0076054.

XX 26-FEB-1998; 98US-0076057.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Olsen HS, Florence K, Brewer LA, Ebner R, Ruben SM, Rosen CA;
 PI Duan RD;

DR N-PSDB; AA210662.

PT WPI: 1999-550857/46.

XX New human genes and the secreted polypeptides they encode, useful for
 PT diagnosis and treatment of e.g. cancers, neurological disorders, immune
 PT diseases, inflammation or blood disorders

XX Claim 11; Page 201; 246pp; English.

XX AAY30701-37 represent human secreted proteins. The polynucleotides and
 CC their corresponding secreted polypeptides are useful for preventing,
 CC therapy or ameliorating medical conditions, e.g. by protein or gene
 CC treatment. Pathological conditions can also be diagnosed by determining
 CC the amount of the new polypeptides in a sample or by determining the
 CC presence of mutations in the polynucleotide. Specific uses include
 CC developing products for the diagnosis or treatment of cancer, tumors,
 CC neurodegenerative disorders, developmental abnormalities and fetal
 CC deficiencies, blood disorders, sepsis, diseases of the immune
 CC system, autoimmune diseases, inflammation, allergies, Alzheimer's and

CC cognitive disorders, schizophrenia, obesity, osteoporosis, arthritis,
 CC infections, AIDS, connective tissue disorders, transplant rejection,
 CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
 CC and reproductive disorders. The polypeptides or polynucleotides can
 CC also be used as food additives or preservatives, such as to increase
 CC or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrate, vitamins, minerals, cofactors or other nutritional
 CC components.

XX Sequence 182 AA;

AAY30723 Length: 182 May 30, 2002 09:25 Type: P Check: 7459 ..

1 MTSRRSSTLS MTSSLLSLGC ALTSAPAST MSWPDLQML DQSPRYMRK

51 SVSOLCPILR PHEPLSSKHP IVLPLQLPPT FLHLLPFGC PGQTVAYWLL

101 EFLSRATLKL YPDGRPLMLQ PRLNFKDHW TTFVSASAL FCVHRMATDR

151 HASEPTHWKA HROGERGHR COHCYSKDL KX

IIAA_SEQUENCE 1.0

ID AAY31622 standard; Protein; 284 AA.

XX AAY31622;

DT 02-NOV-1999 (first entry)

XX Human oxidised LDL receptor (HOLR) CDNA.

DE Human oxidised LDL receptor; HOLR; low density lipoprotein; LDL;
 XX cardiovascular disorder; atherosclerosis; hypertension; stroke;
 XX cancer; autoimmune disorder; diagnosis; treatment.

XX Homo sapiens.

XX Key

XX Modified-site

FT /note= "O-phosphorylated by protein kinase C"

FT /note= "O-phosphorylated by protein kinase C"

FT /note= "O-phosphorylated by protein kinase C"

FT /note= "transmembrane region"

FT /note= "C-type lectin domain"

FT /note= "X encoded by NNN"

FT /note= "O-phosphorylated by casein kinase II"

FT /note= "O-phosphorylated by casein kinase II"

FT /note= "N-glycosylated"

FT /note= "O-phosphorylated by casein kinase II"

FT /note= "O-phosphorylated by tyrosine kinase"

FT /note= "O-phosphorylated by casein kinase II"

FT /note= "O-phosphorylated by tyrosine kinase"

FT /note= "N-glycosylated"

FT /note= "O-phosphorylated by casein kinase II"

FT /note= "O-phosphorylated by protein kinase C"

FT /note= "O-phosphorylated by protein kinase C"

FT /note= "O-phosphorylated by casein kinase II"

PN US5945308-A.

XX 31-AUG-1999.
 PD
 XX
 XX 03-APR-1998; 98US-0055095.
 PF
 XX
 PR 03-APR-1998; 98US-0055095.
 PA
 XX (INCY-) INCYTE PHARM INC.
 PI
 XX Corley NC, Patterson C, Sather S, Tang YF;
 DR WPI: 1999-508172/42.
 DR N-PSDB: AA211726.
 PT
 XX Isolated and purified polynucleotide encoding an LDL receptor
 PS Claim 1; Fig 1; 31pp; English.
 XX
 CC This sequence represents a human oxidised LDL receptor (HOLR).
 CC cDNA sequences encoding HOLR were first identified in a clone from a
 CC foetal lung cDNA library, and a consensus sequence produced from several
 CC overlapping and/or extended nucleic acid sequences. The cDNA, and hence
 CC this sequence are therefore consensuses. High plasma concentrations of
 CC LDL and low plasma concentrations of HDL (high density lipoprotein) are
 CC associated with atherosclerosis and hypertension. Some of the
 CC deleterious effects of LDL may be associated with an oxidised
 CC of LDL, and its uptake and subsequent degradation in vascular
 CC endothelial cells as been implicated in the genesis of atherosclerotic
 CC plaques. The HOLR cDNA and sequences complementary to it may be used
 CC for the diagnosis and/or treatment (antisense therapy) of disorders
 CC associated with the expression of HOLR, including cardiovascular
 CC disorders (e.g., atherosclerosis, hypertension, stroke), various cancers
 CC and autoimmune disorders (e.g., AIDS, diabetes mellitus, multiple
 CC sclerosis, rheumatoid arthritis). HOLR protein antagonists may also be
 CC used to treat such disorders.
 SQ
 XX Sequence 284 AA:
 AAY31622 Length: 284 May 30, 2002 09:25 Type: P Check: 7787 ..
 1 MLDDDDDTTM SLHSQASATA RHPRPRTEH RAPSTWRPV ALTLILCLV
 51 LILGLALGL LCKSALXPEG GSWFOVROY QLSNTGDDTI SOMEERLGNP
 101 SOELQSLQVQ NIKLAGSLQH VAEKLCRELY NKAGAHRCSP CTEOWMKMGD
 151 NCYQFYKDSK SWEDCKYFCL SENSFMKLKIN KQDLEPRAAS QSYSEFFYSY
 201 WTGLLRPDSG KAWLMMDGTP FTSELPHIIT DVTSPRSRCQ VAILNGMIFS
 251 KCKELKRCV CERRAGWKP ESLHVPETL GEGD
 !!AA-SEQUENCE 1.0
 ID AAY38483 standard; Protein: 86 AA.
 XX
 AC AAY38483;
 XX
 DT 18-OCT-1999 (first entry)
 XX
 DE Human secreted protein encoded by gene No. 24.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatitis; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 PN WO9935158-A1.

XX 15-JUL-1999.
 PD
 XX
 XX 06-JAN-1999; 99WO-US00108.
 PF
 XX
 PR 07-JAN-1998; 98US-0070704.
 PA
 XX 07-JAN-1998; 98US-0070657.
 PR 07-JAN-1998; 98US-0070658.
 PR 07-JAN-1998; 98US-0070692.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI
 XX Brewer LA, Duan RD, Ebner R, Lafleur DM, Ni J;
 PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;
 XX
 DR WPI: 1999-444190/37.
 DR N-PSDB: AA206242.
 PT
 XX New isolated human genes and the secreted polypeptides they encode
 PS Disclosure; Page 219; 227pp; English.
 XX
 CC This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. AA206210) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 36 novel genes and their fragments (nucleic
 CC acid sequences: AA206219-206263; amino acid sequences AAY38386-Y38498)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 36
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AA206219 for described uses).
 SQ
 XX Sequence 86 AA:
 AAY38483 Length: 86 May 30, 2002 09:25 Type: P Check: 9545 ..
 1 MRQLFYRDE CGLMKKGGLY FSDFWNKLDV GAILFYAGL TCRILPATLY
 51 PGRVILSLDF ILFCRLMHI FTISKTLGPK IIVYK
 !!AA-SEQUENCE 1.0
 ID AAY38484 standard; Protein: 27 AA.
 XX
 AC AAY38484;
 XX
 DT 18-OCT-1999 (first entry)
 XX
 DE Human secreted protein encoded by gene No. 24.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatitis; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 PN WO9935158-A1.
 XX
 PD 15-JUL-1999.
 XX
 PF 06-JAN-1999; 99WO-US00108.
 XX
 PR 07-JAN-1998; 98US-0070704.
 PA
 XX 07-JAN-1998; 98US-0070657.
 PR 07-JAN-1998; 98US-0070658.
 PR 07-JAN-1998; 98US-0070692.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI
 XX Brewer LA, Duan RD, Ebner R, Lafleur DM, Ni J;
 PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;
 XX
 DR WPI: 1999-444190/37.
 DR N-PSDB: AA206242.
 PT
 XX New isolated human genes and the secreted polypeptides they encode
 PS Disclosure; Page 219; 227pp; English.
 XX
 CC This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. AA206210) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 36 novel genes and their fragments (nucleic
 CC acid sequences: AA206219-206263; amino acid sequences AAY38386-Y38498)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 36
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AA206219 for described uses).
 SQ
 XX Sequence 86 AA:
 AAY38483 Length: 86 May 30, 2002 09:25 Type: P Check: 9545 ..
 1 MRQLFYRDE CGLMKKGGLY FSDFWNKLDV GAILFYAGL TCRILPATLY
 51 PGRVILSLDF ILFCRLMHI FTISKTLGPK IIVYK
 !!AA-SEQUENCE 1.0
 ID AAY38484 standard; Protein: 27 AA.
 XX
 AC AAY38484;
 XX
 DT 18-OCT-1999 (first entry)
 XX
 DE Human secreted protein encoded by gene No. 24.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatitis; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 PN WO9935158-A1.
 XX
 PD 15-JUL-1999.
 XX
 PF 06-JAN-1999; 99WO-US00108.
 XX
 PR 07-JAN-1998; 98US-0070704.
 PA
 XX 07-JAN-1998; 98US-0070657.
 PR 07-JAN-1998; 98US-0070658.
 PR 07-JAN-1998; 98US-0070692.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI
 XX Brewer LA, Duan RD, Ebner R, Lafleur DM, Ni J;
 PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;
 XX
 DR WPI: 1999-444190/37.
 DR N-PSDB: AA206242.
 PT
 XX New isolated human genes and the secreted polypeptides they encode
 PS Disclosure; Page 219; 227pp; English.
 XX
 CC This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. AA206210) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 36 novel genes and their fragments (nucleic
 CC acid sequences: AA206219-206263; amino acid sequences AAY38386-Y38498)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 36
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AA206219 for described uses).
 SQ
 XX Sequence 86 AA:
 AAY38483 Length: 86 May 30, 2002 09:25 Type: P Check: 9545 ..
 1 MRQLFYRDE CGLMKKGGLY FSDFWNKLDV GAILFYAGL TCRILPATLY
 51 PGRVILSLDF ILFCRLMHI FTISKTLGPK IIVYK
 !!AA-SEQUENCE 1.0
 ID AAY38484 standard; Protein: 27 AA.
 XX
 AC AAY38484;
 XX
 DT 18-OCT-1999 (first entry)
 XX
 DE Human secreted protein encoded by gene No. 24.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatitis; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 PN WO9935158-A1.

PR 07-JAN-1998; 98US-0070658.
PR 07-JAN-1998; 98US-0070692.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Brewer LA, Duan RD, Ebnner R, Lafleur DW, Ni J;
XX Olsen HS, Rosen CA, Ruben SM, Soppet DR;
DR WPI: 1999-444190/37.
DR N-PSDB: AA206242.
XX
PT New isolated human genes and the secreted polypeptides they encode
XX
PS Disclosure: Page 219, 227pp: English.
XX
CC This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AA06210) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 36 novel genes and their fragments (nucleic
CC acid sequences: AA206219-206263; amino acid sequences AAY38386-Y38498)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 36
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AA206219 for described uses).
XX
SQ Sequence 27 AA:
AAY38484 Length: 27 May 30, 2002 09:25 Type: P Check: 8317 ..
1 DECGLMKKGG LYFSDFMNKL DVGAILL
!!AA_SEQUENCE 1.0
ID AAY36885 standard; Protein; 945 AA.
XX
AC AAY36885;
XX
DT 07-OCT-1999 (first entry)
XX
DE Protein involved in intermediate metabolism of polypeptides.
XX
KW Vaccine: eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perlepatitis;
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW Bartholinitis; pneumonia; venereal lymphogranulomatosis.
XX
OS Chlamydia trachomatis.
XX
PN WO9928475-A2.
XX
PD 10-JUN-1999.
XX
PE 27-NOV-1998; 98WO-IB01939.
XX
PR 04-NOV-1998; 98US-0107077.
PR 28-NOV-1997; 97FR-0015041.
PR 17-DEC-1997; 97FR-0016034.
XX
PA (GEST) GENSET.
XX
PI Griffiths R;
XX
DR WPI: 1999-371125/31.
XX
PT Genome sequence of Chlamydia trachomatis
XX
PS Disclosure: Page 752-754; 1755pp: English.
XX
CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome

CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
CC perlepatitis, Bartholinitis; pneumonia; venereal lymphogranulomatosis.
CC and venereal lymphogranulomatosis. The polypeptides of the invention
CC may be of use in treating these diseases.
XX
SQ Sequence 945 AA:
AAY36885 Length: 945 May 30, 2002 09:25 Type: P Check: 7481 ..
1 TRIALKEAVL LQRYELIRLI GKGMGEVYL AHDKACRRV ALKRIREDLS
51 GNALLKKRRL REAKIADLI HPGLVPIYYSI CSQGEAVYIT MYIEGFSIK
101 SLKSWQRE VLKLELEERT SVKSLPIFD KICATVEYIH SKGVLRDLIK
151 PDNILLGLRG EVVIVDWGAA IFKAKELKL EQDDEAVSF DERNICYSM
201 TIPGKIVGTP DYMAPESLIG VEASEKDIY ALGLILYQML TLAPPYRRKK
251 GKRLSTRDVV LPPIEMSPYR EIPPSLSQIA MKAIAINPAD RRSIQELRQ
301 ALQPIYQGPB EWTYKATLMA KEKSCWKYYD PILSRYPV LASSPAQWYN
351 FMSEVEISA STRVEYITK SAVHEGMIL FLPSKEKERG ERYCYGLMF
401 SVONHELTVS LKNGIETQK KSQEMISOQY RPAILIEKSD NRIAVFEQA
451 LFLIHIDYLP SLGNRLGVII QDLQMSNIA ISESIGALRV SCLAVDAFL
501 SEKLYDQAAI FYRKIRDSFP GKRESYEQF RLQVTLITQI EEOGGDLTQA
551 LSSFDYLRGG AGAPLEYLQK ALVYQNGSF VEBIRCLLPA LKRYQHPET
601 PRLIEDLCFR LYDSLKHRS EALVFMILLI WIAPEKISVR EEKREPLIY
651 HKQOATLFCQ VDKAPLOFRS SKMELFLSFW TALSLELPEL FRAAGELRY
701 QALADIFVYA GYSGNREAPM QESTALANVS DETTPESLH NOKVAELMF
751 VGVGVALRNK DYQAKAKAFM GKTPTTQLY ALDIFHIQAF IDEETIESFD
801 LLAGIYDPAS EERPHILYV IIOGHMNRD LERAYKLLND RPLDEBLAE
851 YSEAFILMGC YLALTGDRYA VKAHFSRCRY KYGKSALIGK CVDGIDFYL
901 DNLVMEKKM TLFOSYFLLR CLNESPARYE KYQAVLSME NNFFD
!!AA_SEQUENCE 1.0
ID AAY38424 standard; Protein; 509 AA.
XX
AC AAY38424;
XX
DT 30-SEP-1999 (first entry)
XX
DE Human secreted protein.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.

PN WO935158-A1.
XX
PD 15-JUL-1999.
XX
PE 06-JAN-1999; 99WO-US00108.
XX
PR 07-JAN-1998; 98US-0070704.
PR 07-JAN-1998; 98US-0070657.
PR 07-JAN-1998; 98US-0070658.
PR 07-JAN-1998; 98US-0070692.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Brewer LA, Duan RD, Ebner R, Lafleur DM, Ni J;
PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;
DR WPI; 1999-444190/37.
DR N-PSDB; AA206257.
XX
PT New isolated human genes and the secreted polypeptides they encode
XX
PS Claim 11, Page 198-199; 227pp; English.
XX
CC This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AA206210) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 36 novel genes and their fragments (nucleic
CC acid sequences: AA206219-206263; amino acid sequences AAY38498-Y38498)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 36
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AA206219 for described uses).
XX
SQ Sequence 509 AA;
AAY38424 Length: 509 May 30, 2002 09:25 Type: P Check: 463 ..
1 MLAFPLLTG LISFREKRLQ DVGTAPARAR AFTAPVVF HLNTLSTFAF
51 LCLFAYVLNV DFQVPVSWCE CAIYLMFSL VCEEMRQLFY DPDECGIMKK
101 AALYFSDFMN KLDVGAILLF VAGLTGRLRP ATLTPGRVIL SLDTILCLRL
151 LMHTFTSKT LGRPIIIVKR MKKDVFFELF LLAWVVSFG VAKQAILIHN
201 ERRVDMLEFG AVYHSYLTIP GOIPGYIDGV NFNPCHSPN GTDPRKPCP
251 ESDATQGRPA FPEMLTVILL CLYLFTNTIL LNLILIAMFN YTPQVQEHNT
301 DOJMKFORHD LKEEYHGRPA APPETILSH LQLFTKRVVL KTPAKRHKQL
351 KNKLEKNEEA ALTSWEIYKL ENYFQNRQFP QKQREPKIE DISAKVDAMV
401 DLDDLDPLKR SGSMQRIKAS LEEQVAGTAR ALHWIVRTLR ASGSSSEADV
451 PTLASQKAAE EPDAERGRK KTEEPGDSYH VNAHLLYPN CPVTRPCPPT
501 RRCPRGRSS
11AA-SEQUENCE 1.0
ID AAY34694 standard; Protein: 945 AA.
XX
AC AAY34694;
XX
DT 13-SEP-1999 (first entry)
XX
DE C. pneumoniae protein involved in metabolism of polypeptides.
XX

KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.
XX
OS Chlamydia pneumoniae.
XX
PN WO9927105-A2.
XX
PD 03-JUN-1999.
XX
PE 20-NOV-1998; 98WO-1B01890.
XX
PR 04-NOV-1998; 98US-0107078.
PR 21-NOV-1997; 97FR-0014673.
XX
PA (GEST) GENSET.
XX
PI Griffiths R;
XX
DR WPI; 1999-357842/30.
XX
PT Genome sequence of Chlamydia pneumoniae
XX
PS Page 692-694; Disclosure; 1912pp; English.
XX
CC AAY34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
XX
SQ Sequence 945 AA;
AAY34694 Length: 945 May 30, 2002 09:25 Type: P Check: 6409 ..
1 ACTVCLDRD QRSLERYDIV RITGKGNGE VYLAIDPCS RKVALKRIRE
51 DLAEPLLRK RFLREARIAA DLHPGVVPV YTIYSEKDPV YUTMPYIEGY
101 TLKTLKSWM QKESLSKELA EKTSGAGFIS IPHKICPTIE YVHSGILHR
151 DLRPNILLG LFEAVYILDM GAAVACGEEE DLDDIVSKE EYLSRRRTIP
201 GRIVGTPDYM APERILGHPA SKSTDIYALG VVLYOMLTLS FPYRRKKGKK
251 IYLDGQRIPS PQEVAPYREI PPFLSAVVMK MLAVDPQERY SSVYELKEDI
301 ESHLKGSPRW TLTTALPPKK SSSWKLINEPI LLSKTFPMIE VSPASWYSLA
351 ISNIESFSEM RLEYTLSSKG LNEGFGILRP TSENALGGDF YQGYGFWLHI
401 KERTLSVSLV KNSLEIQRS QDLESDKETP LIALQHNHS LSLRPDGTW
451 LIHMYLPSR SGRVAILYRD MEDILEDIGI FESSSGSLKYS CLAVPDAPLA
501 EKLVDRLVL YRRIAESFPG RKEGYEARER AGITVLEKAS TDNNDQERAL
551 AIEESKLHD GVAAPLEYLG KALVYQRLQE YNEEIKSLLL ALKRRYSQHP
601 IFRKDDHYV RLHESFYKRD RLALVMIYV LEIAPQAITP GOEKKILVWL
651 KDKSRATLFC LDDPTVLELR SSKMELFLSY WSGFIPHLNS LFRHAWQSD
701 VRLAIEIFYV ACDLHKWQFL SSCIDIFKES LEDQKATEEI VEFSEFDGA
751 PLFAIQSIFN KEDAEKIFVS NDQLSPILLY YTFDLFANRA LLESQGEAIF
XX

801 QALDLIRSKV PENFYHDYR NHEIRAHLMC RNEKALSTIF ENTERQOKD
 ID EOHLEFVLYG CYALIQGAE AAKOHFVYCR EDRIFPASL ARNYNRGLGP
 851 KDLISYQERR LILROKFLYE HCLGNHDERD LCQTMWHLTF EERQL
 901
 IIAA_SEQUENCE 1.0
 ID AAY05317 standard; Protein; 280 AA.
 XX AAY05317;
 AC
 XX
 DT 25-JUN-1999 (first entry)
 XX
 DE Human secreted protein bn97_1.
 XX
 KW Secreted protein; human; nutritional activity; cytokine; tissue growth;
 KW cell proliferation; cell differentiation; immune stimulant; chemotaxis;
 KW immune suppressant; haematopoiesis regulator; activin; inhibitor; cadherin;
 KW chemokinesis; haemostasis; thrombolysis; anti-inflammatory; gene therapy;
 KW tumour invasion suppressor; tumour inhibitor.
 KM
 OS Homo sapiens.
 XX
 PN MO9913066-AI.
 XX
 PD 18-MAR-1999.
 XX
 PF 08-SEP-1998; 98WO-US18724.
 XX
 PR 08-SEP-1997; 97US-0929007.
 XX
 PA (GEMV) GENETICS INST INC.
 XX
 PI Agostino MJ, Evans C, Jacobs K, Lavallie ER, McCoy JM;
 PI Merberg D, Racie LA, Spaulding V, Treacy M;
 DR WPI: 1999-229235/19.
 DR N-PSDB; AAX33810.
 XX
 PT New polynucleotides encoding secreted human proteins
 XX
 PS Claim 8; Page 79; 96pp; English.
 XX
 CC This sequence is a human secreted protein of the invention. The
 CC secreted proteins were obtained from human adult placenta, foetal brain,
 CC adult testes or adult blood cDNA libraries. The polynucleotides (PNS) and
 CC proteins are predicted to have biological activities which would make
 CC them suitable for treating, preventing or ameliorating medical conditions
 CC in humans and animals, although no supporting data is given. Suggested
 CC activities include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
 CC invasion suppressor activity, and tumour inhibition activity. The PNS are
 CC also stated to be useful for gene therapy.
 XX
 SQ Sequence 280 AA:
 AAY05317 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..

1 MQAKYSTRD MDDDDGDTM SLHSQASAT RHPRPRTEH RAPSSTWRPV
 51 ALTLTLTLCLV LILGLAALGL LFFOYQOLSN TGOPTISOME ERLGNTSOEL
 101 OSLOVQNIKL AGSLQHVAKR LCRELKYNKAG AHRCSPTQD WKHGNCTQ
 151 FYDQSKSWED CKYFCLSENS TMLKINKOED LEPAASOSYS EFFYSTWTGL
 201 LRPDSCRAWL WMDGTPPTSE LFHIIIDVTS PRSRDCAVAIL NCMIFSKDK
 251 ELKRCVCERR AGMWKPESLH VPPTLGECD

IIAA_SEQUENCE 1.0
 ID AAM67815 standard; Protein; 57 AA.
 XX
 AC AAM67815;
 XX
 DT 25-MAR-1999 (first entry)
 XX
 DE Human secreted protein encoded by gene 9 clone HCEEP50.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammatory; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 KM
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 10 /label= unknown
 FT Misc-difference 28 /label= unknown
 FT
 XX
 PN MO9842738-AI.
 XX
 PD 01-OCT-1998.
 XX
 PF 19-MAR-1998; 98WO-US05311.
 XX
 PR 30-MAY-1997; 97US-0050937.
 PR 21-MAR-1997; 97US-0041276.
 PR 21-MAR-1997; 97US-0041277.
 PR 21-MAR-1997; 97US-0041281.
 PR 21-MAR-1997; 97US-0042344.
 PR 30-MAY-1997; 97US-0048069.
 PR 30-MAY-1997; 97US-0048094.
 PR 30-MAY-1997; 97US-0048095.
 PR 30-MAY-1997; 97US-0048096.
 PR 30-MAY-1997; 97US-0048099.
 PR 30-MAY-1997; 97US-0048131.
 PR 30-MAY-1997; 97US-0048135.
 PR 30-MAY-1997; 97US-0048154.
 PR 30-MAY-1997; 97US-0048160.
 PR 30-MAY-1997; 97US-0048186.
 PR 30-MAY-1997; 97US-0048187.
 PR 30-MAY-1997; 97US-0048188.
 PR 30-MAY-1997; 97US-0048350.
 PR 30-MAY-1997; 97US-0048351.
 PR 30-MAY-1997; 97US-0048352.
 PR 30-MAY-1997; 97US-0048355.
 PR 05-AUG-1997; 97US-0054804.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Brewer LA, Duan R, Ebner R, Ferrie AM, Florence KA;
 PI Greene JM, Hu JS, Lafleur DW, Moore PA, Ni J, Olsen HS;
 PI Rosen CA, Ruben SM, Shi Y, Young P;
 DR WPI: 1999-070066/06.
 DR N-PSDB; AAX00619.
 XX
 PT New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 XX
 PS Claim 11; Page 277; 385pp; English.
 XX
 CC This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate

CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. AAX00602) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 87 novel genes and their fragments (nucleic
 CC acid sequences: AAX00611-X00724; amino acid sequences AAW67807-W68004)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 87
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAX00611 for described uses).

XX
 XX
 SO Sequence 57 AA;
 AAW67815 Length: 57 May 30, 2002 09:25 Type: P Check: 5177 ..

1 MLVYDLYLX PKLMALATPO KNGKGRXGD GTPAQAFTWDF WSHLISADPO
 51 TWERAAP
 11AA-SEQUENCE 1.0
 ID AAB56913 standard; Protein: 275 AA.
 XX
 AC AAB56913;
 XX
 XX
 D7 13-MAR-2001 (first entry)
 XX
 DE Human prostate cancer antigen protein sequence SEQ ID NO:1491.
 XX
 KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
 KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200055174-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05988.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 PI Rosen CA, Ruben SM;
 PI
 DR MPI: 2000-587513/55.
 DR N-PSDB; AAF16116.
 XX
 PT Prostate cancer associated gene sequences, referred to as prostate
 PT cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -
 XX
 XX
 PS Claim 11; Page 1927-1928; 2338bp; English.
 XX
 CC AAF15566 to AAF16505 encode the human prostate cancer associated
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytostatic,
 CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
 CC nephrotropic, antiinfective, gynecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotide sequences may be used for detection of prostate cancer,
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to

CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.
 CC
 XX
 XX
 SO Sequence 275 AA;
 AAB56913 Length: 275 May 30, 2002 09:25 Type: P Check: 7508 ..

1 KPEKGVHLN SLDPOQHLM IPLCAPNSLS QLPITDTIRK DSKKKRKA
 51 SKLLMKTYH GMLXSVTEG ASARKTQIPA AQPPRPVPSQ ARPPPNQKGS
 101 SRPPIIIIPA ATTSLITMLN AKDLLQDLKF VPSDEKKKQG QORENETLIQ
 151 RRKDQMPQG TAISVTVPYR VVDQPLKIMP QDMRVVAVF VQGPANQFKG
 201 WPTLLPQSGP VDFIFAKIAF HIKYDEVRLD PNVQKMDVTV LEISYKRRHL
 251 DRPYFLRFWE TLDRYVFKRH SHLRF

11AA-SEQUENCE 1.0
 ID AAB58349 standard; Protein: 100 AA.
 XX
 AC AAB58349;
 XX
 XX
 D7 14-MAR-2001 (first entry)
 XX
 DE Lung cancer associated polypeptide sequence SEQ ID 687.
 XX
 KW Human; lung cancer associated protein; neuroprotective; cytostatic;
 KW cardioactive; immunomodulatory; muscular active; vulnary;
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200055180-A2.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05918.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 PI Ruben SM;
 PI
 DR MPI: 2000-587514/55.
 DR N-PSDB; AAF18225.
 XX
 PT Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer -
 XX
 XX
 PS Claim 11; Page 1192; 1425bp; English.
 XX
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective, cytostatic, cardioactive;
 CC immunomodulatory; muscular active general; vulnary; gastrointestinal
 CC general; nephrotropic; antiinfective; gynecological; of antibacterial
 CC activity. The invention also includes antibodies specific for the
 CC protein or polynucleotide sequences. The lung cancer associated
 CC polynucleotide sequences may be used for detection of lung cancer,
 CC chromosome identification, as chromosome markers, and for numerous other
 CC diagnostic or research purposes. The proteins may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders. The proteins may also be used in the treatment of wounds and
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and

CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterisation of the polynucleotide and protein
CC sequences.

XX Sequence 100 AA;

AAB58349 Length: 100 May 30, 2002 09:25 Type: P Check: 8661 ..

1 STYNTLIPYC OCLFAAPPHF FYIIXTVLIF FCHWDCISDT LHXSLLAIM

51 KSGSKGSGGA XRPGVWXLIO NRNKTPOS LPMSIQLFCC ISCLFEKPIK

11AA_SEQUENCE 1.0

ID AAB03957 standard; Protein: 531 AA.

AC AAB03957;

DT 26-FEB-2001 (first entry)

DE Consensus human mesenchymal stem cell polypeptide.

KW Mesenchymal stem cell; cDNA; detection; genome; marker; locating;
KW genetic marker; linkage analysis; genetic disease; chromosome marker;
KW bone; gene therapy; genetic fingerprinting; peptide; immunogen;
KW antibody; human.

OS Homo sapiens.

PN WO200059933-A2.

PD 12-OCT-2000.

PF 31-MAR-2000; 2000MO-US08751.

PR 01-APR-1999; 990S-0127418.

PR 13-AUG-1999; 990S-0148800.

PA (OSIR-) OSIRIS THERAPEUTICS INC.

PI Van Den Bos C, Mbalaviele G;

DR MPI; 2000-664986/64.

DR N-PSDB; AAB54295, AAB54296.

PT Novel human mesenchymal stem cell polynucleotides and polypeptides
PT useful as reagents for chromosomal mapping, DNA fingerprinting and for
PT detecting presence of genetic mutations in diseases affecting bone
PT growth

PS Claim 16; Page 72-73; 73pp; English.

CC New human mesenchymal stem cell cDNA's and their fragments are useful
CC for detecting genes within a human genome i.e., for detecting the
CC presence of specific mRNA in a particular cell type and in genetic
CC linkage analysis, and for locating gene regions associated with
CC genetic disease, to control gene expression. The cDNA's, fragments
CC and expression products are useful as chromosome markers for
CC determining the location of genes within the genome and alleles
CC expressed, during the development of differentiated mesenchymal cell.
CC The cDNA's are also useful in detecting the presence of mutations
CC in genes leading to diseases affecting bone growth, by comparing the
CC sequences with those found in similar chromosomal locations in
CC mammals, especially humans afflicted with diseases affecting bone
CC growth, in human fingerprinting and in gene therapy. The polypeptides
CC encoded by the mesenchymal stem cell cDNA's or peptide fragments,
CC derivatives, analogs or cells expressing them are useful as
CC immunogens for the production of antibodies.

XX Sequence 531 AA;

AAB03957 Length: 531 May 30, 2002 09:25 Type: P Check: 4914 ..

1 MADVLSVLHQ YNIOKKEIVV KGDEVINGEF SMPKNVKTNY VVWGCRGSGO

11AA_SEQUENCE 1.0
ID AAB36865 standard; Protein: 1503 AA.

AC AAB36865;

DT 21-FEB-2001 (first entry)

DE Human putative Ca2+ channel protein Trpc7 protein.

KW mutTCH-1; sugar-phosphate hydrolase; NUDT9.

OS Homo sapiens.

PN WO200065056-A2.

PD 02-NOV-2000.

PF 26-APR-2000; 2000MO-US11319.

PR 26-APR-1999; 990S-0131051.

PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

PI Scharenberg AM;

DR MPI; 2000-687347/67.

DR N-PSDB; AAC68399.

PT Novel nucleic acids encoding a Mutr domain-containing polypeptide
PT present in human calcium channel protein Trpc7, useful for diagnostic
PT and therapeutic purposes -

PS Claim 39; Page 65; 69pp; English.

CC The present invention relates to human mutTCH-1. Therapeutics
CC involving mutTCH-1 may be useful for diagnosing and treating
CC conditions associated with aberrant levels of expression of (II) and
CC for identifying agents that are useful for treating diseases associated
CC with pyrophosphohydrolase and/or sugar-phosphate hydrolase activity.

XX Sequence 1503 AA;

AAB36865 Length: 1503 May 30, 2002 09:25 Type: P Check: 5055 ..

1 MEPSALRKAG SEQEGFEGFL PRRTYDLAGV SNLRNSNSI FSKWLOQCP

51 GNNDKOESLS SWIPENIKKK ECVYFVSSK LSDAGKVVQO CGYTHQHL

101 EATKPHTFQG TOWDPKXHVQ EMPTAFGDI VFTGLSQXYK KYRVASQDTP

151 SSVIYHLMTQ HMGIDVNPUL ISVTGAKNF NMPRLKSTF RRGLYKVAQT

201 TGAMITGGS HTGVMKQVGE AVRDEFSLSSS YKEGELLITIG VATWGTVHRR
 251 EGLHHPGSE PAVYIIDEDE QGNLTCLODN HSHFLIVDGG THGOYGEIIP
 301 LRTPLEKFIS EOTKERGVA IKIPIVCYL EGGPGLTHTI DNATTNGTPC
 351 VVVEGSGRVA DIVAQVANLP VSDITISLIO OKLSVFPOEM FEFTESRIV
 401 EMTKKIDIV RRROLTVFR ECKDQGDVD VAILQALLKA SRSDHKGHE
 451 NMDHQLKLV AMNRVIDIARS EIFMDEMOMK PSDLHPMTA ALISNKEFV
 501 KLFENGVOYL KEFVTWDTLL YLYENIDPSC LFHSKLOKVL VEDPERPACA
 551 PAARLQMH VAQVLRRELG DFTQPLYP RPNDLRLL PVPRVXLMVQ
 601 GVSRLSLYKR SSGHVTFTMD PIRDLIMAI VONRRELGI IMAOSQOCIA
 651 AALACSKILK ELKREEDTD SSEEMLALAE EYENRAIGVF TECYRKDEER
 701 AOKLTRYSE AMGKTICLOL ALEAKDMKFV SHGIGQAFLT KYWVGQLSVD
 751 NGLRVLICM LAPPILITGL ISFRERKLOD VGTPAARARA FTTPAVVVFH
 801 LNLSTYFAFL CLPFAVYLMVD FQPVPSWCEC AIYLMFLSV CEENROLEFD
 851 PDEGLMKKA ALVPSDFWNK LDVGAILLFV AGLTCRLIPA TLYPGVILS
 901 LDFLFLCLRL MHFTTISKTL GKRIIIVKRM MKDVEFFLFL LAVWVSEGV
 951 AKQAILIINE RAVDWLFRGA VYHSYLTIFG QIPGYIDGVN FNPHCSFNG
 1001 TDPKPKPCE SDATQGRPAF PEMLVFVLLLC LYLLFTNILL LNLIAFENV
 1051 TFOQVQEHDT QIMKFORHDL IEEYHRPAA PPPFLLSHL QLFKRYVVK
 1101 TPARRHKQLK NLEKNEEA LLSWEIYLKE NYLONROFOQ KQREPKTIED
 1151 ISNKVDAMVD LLDLDPLKRS GSWEQRLASL EEOVAQTARA LHWIVRLRA
 1201 SGFSSEADVP TLASQKAAE PDAEPGRKK TEEPDSTHV NARHLIPNC
 1251 PVTRFPVNE KVPWETEFLI YDPPEYTAER KDAAMDPNG DTLEPLSTIQ
 1301 YNVVDGLDR RSPHGPYTVQ AGPLNPMGR TGLRGSLIS CFGNHTLYP
 1351 MVTMRNRNED GAICRSIKK MLEVLYVKLP LSEHWALPGG SREGEMLPR
 1401 KLKRLROEH WPSFENLLKC GMEVYKGYMD DFRNTDNAMI ETYAVSVHFQ
 1451 DONVELMRL NSNLHACDSG ASIRMOVDR RPLVANHKT LLOKAAAEFG
 1501 AHY
 11AA_SEQUENCE 1.0
 ID AAB12123 standard; Protein: 344 AA.
 XX
 AC AAB12123:
 XX
 DT 02-FEB-2001 (first entry)
 XX
 DE Hydrophobic domain protein from clone HP10608 isolated from Saos-2 cells.
 XX
 KW Human; secreted protein; membrane protein; hydrophobic domain;
 KW proliferation control; differentiation induction; material transport;
 KW biophylaxis; signal receptor; ion channel; transporter; immunostimulant;
 KW immunosuppressant; haematopoiesis regulator; chemotactic; chemokine;
 KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
 KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;
 XX
 OS Homo sapiens.
 XX
 FN WO200029448-A2.

XX 25-MAY-2000.
 XX 17-NOV-1999; 99WO-JP06412.
 XX 17-NOV-1998; 98JP-0326255.
 XX 22-DEC-1998; 98JP-0364315.
 XX 16-MAR-1999; 99JP-0069811.
 XX 27-APR-1999; 99JP-0119299.
 XX 19-MAY-1999; 99JP-0138169.
 XX (SAGA) SAGAMI CHEM RES CENT.
 PA (PROT-) PROTEGENE INC.
 XX
 PI Kato S, Kimura T;
 XX
 DR WPI: 2000-387753/33.
 DR N-PSDB: AAA60185, AAA60195.
 XX
 PT Proteins comprising hydrophobic regions, such as secretory and membrane
 PT proteins, useful in research and diagnostics and having various
 PT activities e.g. immunomodulatory, antinflammatory, chemokinetic,
 PT hemostatic, thrombolytic -
 XX
 PS Claim 1; Page 188-189; 410pp; English.
 XX
 CC Secretory proteins play important roles in the proliferation control, the
 CC differentiation induction, the material transport and the biophylaxis of
 CC cells. Membrane proteins have important roles as signal receptors, ion
 CC channels and transporters. The present sequence is a human protein which
 CC has at least one hydrophobic domain. This protein may be a secretory or a
 CC membrane protein. The present protein may have cytokine and cell
 CC proliferation/differentiation activity, immune stimulating or suppressing
 CC activity, haematopoiesis activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic
 CC and thrombolytic activity, anti-inflammatory activity and tumour
 CC inhibition activity. The present protein could therefore be used for
 CC treatment of autoimmune disease, Alzheimer's disease, Parkinson's
 CC disease, and cancer.
 XX
 SO Sequence 344 AA;
 AAB12123 Length: 344 May 30, 2002 09:25 Type: P Check: 1439 ..
 1 MQPPPGPLG DCLRDWEDIQ ODFONIQETH RLYRLKLEEL TKLQNNCTSS
 51 ITRQKKRIQE LALAKKCKP SLPAEABGAA QLELQMKER QGLFDMEMAY
 101 LPKKNGLYLS LVIGNVVTL LSKQAKFAYR DEYEKFKLYL TIIILISFT
 151 CRFLNSRVT DAAVNFLLW YCTLTITRES ILINNGSRK GWWVFHHYVS
 201 TFLSGVMLTW PDGLMYOKFR NOFLSFSMWQ SFVQPLQIYY QSGCLYRLRA
 251 LGERHTMDLT VEGFQSWMMR GLTFLLPFLR FGHFQOLFNA LTFLELADP
 301 CCKEQVLMC GFPLLLFLG NFFTLLRVNH HKFHSORHGS KKDD
 11AA_SEQUENCE 1.0
 ID AAB41885 standard; Protein: 251 AA.
 XX
 AC AAB41885:
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF1649 polypeptide sequence SEQ ID NO:3298.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antiparastic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopthic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;

KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 KM
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PE 31-MAR-2000; 2000MO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR MPI: 2000-602362/57.
 DR N-PSDB; AAC76094.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 2491; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in ABA40237 to ABA43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytosstatic; hepatotropic; vulnerary;
 CC antipsoptic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antifungal; antirheumatic; antihypoid;
 CC antihypoid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX
 SO Sequence 251 AA:

AA041885 Length: 251 May 30, 2002 09:25 Type: P Check: 2227 ..

1 GHLCVCLGVC MYLCVVCCLG VCAALCVCL CECILWVFCV CVCICVCYL
 51 CACMCIDVFC CMKLCVCLIV CICYVVCVCH PCEFWCLSA CLCIPVSPCV
 101 CLCVCICXCL CMCVGCVSV CVCVCIEREG ERKATDGSX WKVYPSQPM
 151 EESVNPPTQD DQIMWCLADS GNVTFILRMG IHLGLEKCS WSLKCEFFP
 201 FVIERAQPCV HMLVTYNLRV GDSHRETEG TADSEQESGG TSLPLGPNQ
 251 L

11AA_SEQUENCE 1.0
 ID ABA42619 standard; Protein: 280 AA.
 XX
 AC ABA42619;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF2383 polypeptide sequence SEQ ID NO:4766.
 XX
 KM Human: open reading frame; ORFX; detection; cytosstatic; hepatotropic;
 KM vulnerary; antipsoptic; antiparkinsonian; nootropic; neuroprotective;
 KM anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KM hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KM antiviral; antibacterial; antifungal; antirheumatic; antihypoid;
 KM antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KM cholesterol ester storage; systemic lupus erythematosus; infection;
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KM bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KM thrombosis; contraceptive.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PE 31-MAR-2000; 2000MO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR MPI: 2000-602362/57.
 DR N-PSDB; AAC76828.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 3947-3948; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in ABA40237 to ABA43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytosstatic; hepatotropic; vulnerary;
 CC antipsoptic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antifungal; antirheumatic;
 CC antihypoid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 XX
 SO Sequence 280 AA:

XX 31-OCT-2000 (first entry)
XX Mdm/hdm antagonist peptide sequence SEQ ID NO:133.
DE
XX
KM Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KM autoimmune disease; cytostatic; antitumour; thrombolytic; VEGF;
KM immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KM MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KM cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KM vascular endothelial growth factor; matrix metalloproteinase;
KM asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
XX 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
XX (AMGE-) AMGEN INC.
PA
PI Feige U, Liu C, Cheetham J, Boone TC;
PI WPI; 2000-350702/30.
DR
XX
XX Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 241; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-R1-(X2)b, where: R1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)-c-P1, -(L1)c-P1-(L2)-d-P2,
CC -(L1)c-P1-(L2)-d-P2-(L3)-e-P3, or -(L1)c-P1-(L2)-d-P2-(L3)-e-P3-(L4)-f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antitumour, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AA69443
CC to AA69526 and AA6955 to AA6955 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;
AAB17077 Length: 12 May 30, 2002 09:25 Type: P Check: 6124 ..
1 QPFFSDYWKI LP
11AA_SEQUENCE 1.0
ID AAB17078 standard; Peptide; 12 AA.
AC AAB17078;
XX
XX 31-OCT-2000 (first entry)
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:134.
XX
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KM autoimmune disease; cytostatic; antitumour; thrombolytic; VEGF;
KM immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KM immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KM

KM MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KM cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KM vascular endothelial growth factor; matrix metalloproteinase;
KM asthma; thrombosis; pharmaceutical.
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
XX 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
XX (AMGE-) AMGEN INC.
PA
PI Feige U, Liu C, Cheetham J, Boone TC;
PI WPI; 2000-350702/30.
DR
XX
XX Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 241; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-R1-(X2)b, where: R1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)-c-P1, -(L1)c-P1-(L2)-d-P2,
CC -(L1)c-P1-(L2)-d-P2-(L3)-e-P3, or -(L1)c-P1-(L2)-d-P2-(L3)-e-P3-(L4)-f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antitumour, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AA69443
CC to AA69526 and AA6955 to AA6955 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;
AAB17078 Length: 12 May 30, 2002 09:25 Type: P Check: 6146 ..
1 QPFFSDYWKI LP
11AA_SEQUENCE 1.0
ID AAB17079 standard; Peptide; 12 AA.
AC AAB17079;
XX
XX 31-OCT-2000 (first entry)
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:135.
XX
XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KM autoimmune disease; cytostatic; antitumour; thrombolytic; VEGF;
KM immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KM MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KM cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KM vascular endothelial growth factor; matrix metalloproteinase;
KM asthma; thrombosis; pharmaceutical.
XX
XX Synthetic.
OS
PN WO200024782-A2.

XX 04-MAY-2000.
 PD 25-OCT-1999; 99WO-US25044.
 XX
 XX 23-OCT-1998; 98US-0105371.
 PR 22-OCT-1999; 99US-0428082.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Feige U, Liu C, Cheetham J, Boone TC;
 XX
 DR WPI; 2000-350702/30.
 XX
 PT Novel composition of matter comprising an Fc domain and
 PT pharmacologically active peptides, useful for treating cancer and
 PT autoimmune diseases -
 XX
 PS Claim 39; Page 241; 608bp; English.
 XX
 CC The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1-(L1)d-P2-(L2)d-P2,
 CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
 CC where P1, P2, P3, and P4 = are each independently sequences of
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
 CC independently linkers; and a, b, c, d, e, and f = are each independently
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
 CC have cytostatic, antitumoral, thrombolytic and immunosuppressive
 CC activities. DNAs, vectors and host cells from the present invention can
 CC be used for producing pharmaceutical compositions. The compositions are
 CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
 CC The use of an Fc domain (rather than a Fab domain) can provide a longer
 CC half-life or incorporate functions such as Fc receptor binding, protein
 CC A binding, complement fixation, and possibly placental transfer. AA69443
 CC to AA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
 CC sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 12 AA;
 XX
 AAB17079 Length: 12 May 30, 2002 09:25 Type: P Check: 5978 ..
 XX
 1 MPRFMDYWG LN
 XX
 !!AA-SEQUENCE 1.0
 ID AAB17080 standard; Peptide; 12 AA.
 XX
 AC AAB17080;
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE Mdm/Idm antagonist peptide sequence SEQ ID NO:136.
 XX
 KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antitumoral; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
 KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase;
 KW asthma; thrombosis; pharmaceutical.
 XX
 OS Synthetic.
 XX
 PN WO200024782-A2.
 XX
 PD 04-MAY-2000.
 XX
 PF 25-OCT-1999; 99WO-US25044.
 XX
 PR 23-OCT-1998; 98US-0105371.
 PR 22-OCT-1999; 99US-0428082.
 XX

PA (AMGE-) AMGEN INC.
 XX
 PI Feige U, Liu C, Cheetham J, Boone TC;
 XX
 DR WPI; 2000-350702/30.
 XX
 PT Novel composition of matter comprising an Fc domain and
 PT pharmacologically active peptides, useful for treating cancer and
 PT autoimmune diseases -
 XX
 PS Claim 39; Page 242; 608bp; English.
 XX
 CC The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1-(L1)d-P2-(L2)d-P2,
 CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
 CC where P1, P2, P3, and P4 = are each independently sequences of
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
 CC independently linkers; and a, b, c, d, e, and f = are each independently
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
 CC have cytostatic, antitumoral, thrombolytic and immunosuppressive
 CC activities. DNAs, vectors and host cells from the present invention can
 CC be used for producing pharmaceutical compositions. The compositions are
 CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
 CC The use of an Fc domain (rather than a Fab domain) can provide a longer
 CC half-life or incorporate functions such as Fc receptor binding, protein
 CC A binding, complement fixation, and possibly placental transfer. AA69443
 CC to AA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
 CC sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 12 AA;
 XX
 AAB17080 Length: 12 May 30, 2002 09:25 Type: P Check: 6151 ..
 XX
 1 VQNFIDYWTQ QF
 XX
 !!AA-SEQUENCE 1.0
 ID AAB17081 standard; Peptide; 12 AA.
 XX
 AC AAB17081;
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE Mdm/Idm antagonist peptide sequence SEQ ID NO:137.
 XX
 KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antitumoral; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
 KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase;
 KW asthma; thrombosis; pharmaceutical.
 XX
 OS Synthetic.
 XX
 PN WO200024782-A2.
 XX
 PD 04-MAY-2000.
 XX
 PF 25-OCT-1999; 99WO-US25044.
 XX
 PR 23-OCT-1998; 98US-0105371.
 PR 22-OCT-1999; 99US-0428082.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Feige U, Liu C, Cheetham J, Boone TC;
 XX
 DR WPI; 2000-350702/30.
 XX
 PT Novel composition of matter comprising an Fc domain and
 PT pharmacologically active peptides, useful for treating cancer and

PT autoimmune diseases -
 XX
 XX Claim 39; Page 242; 608pp; English.
 XX
 CC The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers, where (I) is:
 CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
 CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
 CC where P1, P2, P3, and P4 = are each independently sequences of
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
 CC independently linkers; and a, b, c, d, e, and f = are each independently
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
 CC have cytostatic, antitasthmatic, thrombolytic and immunosuppressive
 CC activities. DNAs, vectors and host cells from the present invention can
 CC be used for producing pharmaceutical compositions. The compositions are
 CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
 CC The use of an Fc domain (rather than a Fab domain) can provide a longer
 CC half-life or incorporate functions such as Fc receptor binding, protein
 CC A binding, complement fixation, and possibly placental transfer. AA69443
 CC to AA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
 CC sequences used in the exemplification of the present invention.
 CC
 SQ Sequence 12 AA;
 AAB17081 Length: 12 May 30, 2002 09:25 Type: P Check: 5993 ..
 1 TGPATHTYMA TF
 11AA_SEQUENCE 1.0
 ID AAB17082 standard; Peptide; 15 AA.
 AC
 AC AAB17082;
 DT 31-OCT-2000 (first entry)
 DE
 DE Mdm/hdm antagonist peptide sequence SEQ ID NO:138.
 XX
 KM Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KM autoimmune disease; cytostatic; antitasthmatic; thrombolytic; VEGF;
 KM immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
 KM MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KM cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KM vascular endothelial growth factor; matrix metalloproteinase;
 KM asthma; thrombosis; pharmaceutical.
 XX
 OS Synthetic.
 OS
 OS WO200024782-A2.
 PN
 PN 04-MAY-2000.
 PD
 PD 25-OCT-1999; 99WO-US25044.
 PF
 PF 23-OCT-1998; 98US-0105371.
 PR 22-OCT-1999; 99US-0428082.
 XX
 XX (AMGE-) AMGEN INC.
 PA
 PA Felge U, Liu C, Cheetham J, Boone TC;
 PI
 PI WPI: 2000-350702/30.
 DR
 DR
 XX Novel composition of matter comprising an Fc domain and
 XX pharmacologically active peptides, useful for treating cancer and
 XX autoimmune diseases -
 XX
 PS Claim 39; Page 242; 608pp; English.
 CC
 CC The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers, where (I) is:
 CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,

CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
 CC where P1, P2, P3, and P4 = are each independently sequences of
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
 CC independently linkers; and a, b, c, d, e, and f = are each independent
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
 CC have cytostatic, antitasthmatic, thrombolytic and immunosuppressive
 CC activities. DNAs, vectors and host cells from the present invention can
 CC be used for producing pharmaceutical compositions. The compositions are
 CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
 CC The use of an Fc domain (rather than a Fab domain) can provide a longer
 CC half-life or incorporate functions such as Fc receptor binding, protein
 CC A binding, complement fixation, and possibly placental transfer. AA69443
 CC to AA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
 CC sequences used in the exemplification of the present invention.
 CC
 SQ Sequence 15 AA;
 AAB17082 Length: 15 May 30, 2002 09:25 Type: P Check: 9093 ..
 1 IDRAPFPRDH WFLV
 11AA_SEQUENCE 1.0
 ID AAB17083 standard; Peptide; 15 AA.
 AC
 AC AAB17083;
 DT 31-OCT-2000 (first entry)
 DE
 DE Mdm/hdm antagonist peptide sequence SEQ ID NO:139.
 XX
 KM Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KM autoimmune disease; cytostatic; antitasthmatic; thrombolytic; VEGF;
 KM immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
 KM MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KM cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KM vascular endothelial growth factor; matrix metalloproteinase;
 KM asthma; thrombosis; pharmaceutical.
 XX
 OS Synthetic.
 OS
 OS WO200024782-A2.
 PN
 PN 04-MAY-2000.
 PD
 PD 25-OCT-1999; 99WO-US25044.
 PF
 PF 23-OCT-1998; 98US-0105371.
 PR 22-OCT-1999; 99US-0428082.
 XX
 XX (AMGE-) AMGEN INC.
 PA
 PA Felge U, Liu C, Cheetham J, Boone TC;
 PI
 PI WPI: 2000-350702/30.
 DR
 DR
 XX Novel composition of matter comprising an Fc domain and
 XX pharmacologically active peptides, useful for treating cancer and
 XX autoimmune diseases -
 XX
 PS Claim 39; Page 243; 608pp; English.
 CC
 CC The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers, where (I) is:
 CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
 CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
 CC where P1, P2, P3, and P4 = are each independently sequences of
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
 CC independently linkers; and a, b, c, d, e, and f = are each independent
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
 CC have cytostatic, antitasthmatic, thrombolytic and immunosuppressive
 CC activities. DNAs, vectors and host cells from the present invention can
 CC be used for producing pharmaceutical compositions. The compositions are

CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 15 AA;
AAB17083 Length: 15 May 30, 2002 09:25 Type: P Check: 9428 ..
1 PRAIVEADY WETLY
11AA_SEQUENCE 1.0
ID AAB17084 standard; Peptide; 15 AA.
XX
AC AAB17084;
XX
DT 31-OCT-2000 (first entry)
XX
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:140.
XX
KM Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KM autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KM immunosuppressive; Epo; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KM MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KM cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KM vascular endothelial growth factor; matrix metalloproteinase;
KM asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
XX (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
XX WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
XX
XX Claim 39; Page 243; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)-C-P1, -(L1)-C-P1-(L2)-d-P2,
CC -(L1)-C-P1-(L2)-d-P2-(L3)-e-P3, or -(L1)-C-P1-(L2)-d-P2-(L3)-e-P3-(L4)-f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 15 AA;
AAB17085 Length: 15 May 30, 2002 09:25 Type: P Check: 8833 ..
1 PAFSRFMSDL SAGAH
11AA_SEQUENCE 1.0
ID AAB17085 standard; Peptide; 15 AA.
XX
AC AAB17085;
XX
DT 31-OCT-2000 (first entry)
XX
DE Mdm/hdm antagonist peptide sequence SEQ ID NO:141.
XX
KM Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KM autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KM immunosuppressive; Epo; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KM MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KM cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KM vascular endothelial growth factor; matrix metalloproteinase;
KM asthma; thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US25044.
XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
XX (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
XX WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
XX
XX Claim 39; Page 243; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)-C-P1, -(L1)-C-P1-(L2)-d-P2,
CC -(L1)-C-P1-(L2)-d-P2-(L3)-e-P3, or -(L1)-C-P1-(L2)-d-P2-(L3)-e-P3-(L4)-f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 15 AA;
AAB17086 Length: 15 May 30, 2002 09:25 Type: P Check: 8896 ..
1 PAFSRFMSKTL SAGAH
11AA_SEQUENCE 1.0
ID AAB17086 standard; Peptide; 10 AA.
XX

11AA_SEQUENCE 1.0
ID AAB17086 standard; Peptide; 10 AA.
XX

AC AAB17086;
 DT 31-OCT-2000 (first entry)
 DE Mdm/hdm antagonist peptide sequence SEQ ID NO:142.
 XX
 XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KM autoimmune disease; cytostatic; antilasthmatic; thrombolytic; VEGF;
 KM immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
 KM MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KM cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KM vascular endothelial growth factor; matrix metalloproteinase;
 KM asthma; thrombosis; pharmaceutical.
 XX
 OS Synthetic.
 XX
 PN WO200024782-A2.
 XX
 PD 04-MAY-2000.
 XX
 PF 25-OCT-1999; 99WO-US25044.
 XX
 PR 23-OCT-1998; 98US-0105371.
 XX
 PR 22-OCT-1999; 99US-0428082.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Felge U, Liu C, Cheetham J, Boone TC;
 XX
 DR WPI; 2000-350702/30.
 XX
 PT Novel composition of matter comprising an Fc domain and
 PT pharmacologically active peptides, useful for treating cancer and
 PT autoimmune diseases -
 XX
 PS Claim 39; Page 243; 608pp; English.
 XX
 CC The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
 CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
 CC where P1, P2, P3, and P4 = are each independently sequences of
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
 CC independently linkers; and a, b, c, d, e, and f = are each independently
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
 CC have cytostatic, antilasthmatic, thrombolytic and immunosuppressive
 CC activities. DNAs, vectors and host cells from the present invention can
 CC be used for producing pharmaceutical compositions. The compositions are
 CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
 CC The use of an Fc domain (rather than a Fab domain) can provide a longer
 CC half-life or incorporate functions such as Fc receptor binding, protein
 CC A binding, complement fixation, and possibly placental transfer. AA69443
 CC to AA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
 CC sequences used in the exemplification of the present invention.
 XX
 SO Sequence 10 AA;
 AAB17086 Length: 10 May 30, 2002 09:25 Type: P Check: 4557 ..
 1 PFXFDYXXL
 IIAA_SEQUENCE 1.0
 ID AAB17089 standard; Peptide; 12 AA.
 AC AAB17089;
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE Mdm/hdm antagonist peptide sequence SEQ ID NO:145.
 XX
 KM Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KM autoimmune disease; cytostatic; antilasthmatic; thrombolytic; VEGF;
 KM

KM immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
 KM MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KM cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KM vascular endothelial growth factor; matrix metalloproteinase;
 KM asthma; thrombosis; pharmaceutical.
 XX
 OS Synthetic.
 XX
 PN WO200024782-A2.
 XX
 PD 04-MAY-2000.
 XX
 PF 25-OCT-1999; 99WO-US25044.
 XX
 PR 23-OCT-1998; 98US-0105371.
 XX
 PR 22-OCT-1999; 99US-0428082.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Felge U, Liu C, Cheetham J, Boone TC;
 XX
 DR WPI; 2000-350702/30.
 XX
 PT Novel composition of matter comprising an Fc domain and
 PT pharmacologically active peptides, useful for treating cancer and
 PT autoimmune diseases -
 XX
 PS Claim 39; Page 244; 608pp; English.
 XX
 CC The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
 CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
 CC where P1, P2, P3, and P4 = are each independently sequences of
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
 CC independently linkers; and a, b, c, d, e, and f = are each independently
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
 CC have cytostatic, antilasthmatic, thrombolytic and immunosuppressive
 CC activities. DNAs, vectors and host cells from the present invention can
 CC be used for producing pharmaceutical compositions. The compositions are
 CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
 CC The use of an Fc domain (rather than a Fab domain) can provide a longer
 CC half-life or incorporate functions such as Fc receptor binding, protein
 CC A binding, complement fixation, and possibly placental transfer. AA69443
 CC to AA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
 CC sequences used in the exemplification of the present invention.
 XX
 SO Sequence 12 AA;
 AAB17089 Length: 12 May 30, 2002 09:25 Type: P Check: 6124 ..
 1 QETFSYWKLP
 IIAA_SEQUENCE 1.0
 ID AAB17090 standard; Peptide; 12 AA.
 AC AAB17090;
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE Mdm/hdm antagonist peptide sequence SEQ ID NO:146.
 XX
 KM Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KM autoimmune disease; cytostatic; antilasthmatic; thrombolytic; VEGF;
 KM immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
 KM MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KM cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KM vascular endothelial growth factor; matrix metalloproteinase;
 KM asthma; thrombosis; pharmaceutical.
 XX
 OS Synthetic.
 XX

PN WO200024782-A2.
 XX
 PD 04-MAY-2000.
 XX
 PF 25-OCT-1999; 99WO-US25044.
 XX
 PR 23-OCT-1998; 98US-0105371.
 XX
 PR 22-OCT-1999; 99US-0428082.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Feige U, Liu C, Cheatham J, Boone TC;
 XX
 DR WPI; 2000-350702/30.
 XX
 PT Novel composition of matter comprising an Fc domain and
 PT pharmacologically active peptides, useful for treating cancer and
 PT autoimmune diseases -
 XX
 PS Claim 39; Page 245; 608pp; English.
 XX
 CC The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
 CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
 CC where P1, P2, P3, and P4 = are each independently sequences of
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
 CC independently linkers; and a, b, c, d, e, and f = are each independently
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
 CC have cytostatic, antitumoric, thrombolytic and immunosuppressive
 CC activities. DNAs, vectors and host cells from the present invention can
 CC be used for producing pharmaceutical compositions. The compositions are
 CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
 CC The use of an Fc domain (rather than a Fab domain) can provide a longer
 CC half-life or incorporate functions such as Fc receptor binding, protein
 CC A binding, complement fixation, and possibly placental transfer. AA69443
 CC to AA69526 and AA6955 to AA69003 represent nucleotide and amino acid
 CC sequences used in the exemplification of the present invention.
 XX
 SO Sequence 12 AA;
 AAB17090 Length: 12 May 30, 2002 09:25 Type: P Check: 6146 ..
 1 QPTFSDYWKLP LP
 IIA_SEQUENCE 1.0
 ID AAY92944 standard; Protein: 1503 AA.
 XX
 AC AAY92944;
 XX
 DT 08-NOV-2000 (first entry)
 XX
 DE Human TRPC7 protein.
 XX
 KM Transmembrane protein; TRPC7; brain; transient receptor potential; TRP;
 KM calcium channel function; human; gene therapy; periodic psychosis;
 KM mutation.
 XX
 OS Homo sapiens.
 XX
 PN WO200029571-A1.
 XX
 PD 25-MAY-2000.
 XX
 PF 11-NOV-1999; 99WO-JP06289.
 XX
 PR 12-NOV-1998; 98JP-0321200.
 XX
 PA (EIKE) EIKEN KAGAKU KK.
 XX
 PI Shimizu N, Nagamine K;
 XX

DR WPI; 2000-387784/33.
 DR N-PSDB; AAA11284.
 XX
 PT Nucleic acids encoding transmembrane protein TRPC7 expressed in brain
 PT and homologous to transient receptor potential protein useful in the
 PT treatment of associated diseases such as periodic psychosis -
 XX
 PS Claim 1; Page 64-71; 77pp; Japanese.
 XX
 CC The invention relates to the isolation of a nucleic acid encoding a
 CC transmembrane protein TRPC7 which is expressed in brain and is homologous
 CC to transient receptor potential (TRP) protein. This suggests that the
 CC TRPC7 protein may have a calcium channel function. This sequence
 CC represents the human TRPC7 protein. The DNA and protein can be used
 CC in the diagnosis and treatment of disorders associated with TRPC7,
 CC especially the screening, monitoring and treatment (by gene therapy)
 CC of periodic psychosis, which appears to be associated with mutations
 CC in the TRPC7 gene.
 XX
 SO Sequence 1503 AA;
 AAY92944 Length: 1503 May 30, 2002 09:25 Type: P Check: 5055 ..

1 MEDSALRKAG SEQEGEGEL PRVYDGLV SNLRSSSL EKSRLQCP
 51 GNNDKQSLSS SWIPENIKKK ECVYEVESKK LSDAGKYVCQ CGYTHQHLE
 101 EARKPHFQCG TQMDPKKHVQ EMPTDAPGDI VFTGLSQKV KYRVSGODTP
 151 SSVYIHLMTQ HMGLDVPNLL ISVTGAKNF NMRDLKSIK RGLVRYAQT
 201 TGAVIITGGS HIGVMKQVE AVRDFSLSS YKEGELITIG VATWGTVHR
 251 EGLIHPTGSF PAVYIIDEQG QGNLTCDSN HSHLLIVDDG THQGYVEIP
 301 LTRFLKEFIS EOKTERGVA IKIPIVCVL EGGPGTLHTI DNATNGTPC
 351 VVVGSGSRVA DVAQVANLP VSDITISLIQ QKLSVFQEM FEFTESRIY
 401 EWTKKIIDIY RRRQLLTFR BSKDQGDVD VALIQALLKA SRQDDHFGHE
 451 NMDHQLKLV ANNRDIANS EITFDEQWK PSDLHPTMTA ALISNKEFY
 501 KLFLENGVOL KEFVYWDTLI YLYENLDPSK LFHSKLQKVL VEDPERPACA
 551 PAAPRLQMH VAVYLRBELG DFTQPLPRP RHNRRLNLL PVPHVKLVNQ
 601 GVSLSLSLYR SSGHVTFTMD PIRDLINAI VQNRRELAGI IMAOSODCIA
 651 AALACSKILK ELKSEEDTD SSEEMIALAE EYEHRAIGVF TECYRKDEER
 701 AQKLTFRVSE AMKTCICQL ALEAKDMKFV SHGIGQAFIT KYWAGQLSVD
 751 NGIMRVTLCK LAEPILLTGL ISFERKRLQD VGTAPARARA EFTAPVYVF
 801 LNLISYFAFL CLFAYVLMVD EQVPVSWCEC AIYIMLESVY CEBMROLFYT
 851 PDECGLMKKA ALYFSDPMNK LDVGAILLFV AGLCRLIPA TLVPGRYILS
 901 LDTFLFCLRL MHITTSKTL GPKIIVKRM MKDYFFELFL LAVVVSFQV
 951 AKQAILIHNE RRYDMLFRGA VYHSYLTIFG QIPGYIDGVN FNEBHSPPNG
 1001 TDPYKPKCPE SDATQORAF PEMLVVLLIC LYLFTNILL LNLIMFNY
 1051 TFOOVQEHND QIKKFORHDL IEEYHGRPA PPTIILSHL QLTIKRYVLK
 1101 TPAKRHRQLK NKLKNEEA LLSWEIYKE NYLONROFOQ KQRPEDKIED
 1151 ISNKVDAMD LLDLDPLKRS GSMQRIASL EGYVAQTARA LHMIVTRLRA
 1201 SGFSSEADVP TLASQKAAE PDAEPGRKK TEEPGDSYHV NARHLLYPNC

1251 PVRFPVPNE KYPMETEFLL YDPEFYTAER KDAAMPMPG DTEEPSTIO
1301 YNVVDGLDR RSPHGPIYVQ AGLPLNPGR TGLRGSLG CCGPNHTLYP
1351 MYTRWRNED GAICRKSIRK MLEVLYVKLP LSEHWALPGC SREPGMLPR
1401 KKKRIIRQEH WPSFENLKC GMEYKKGKMD DFRNTDAMI EYVAVSVHQ
1451 DONDVELNRL NSNLHACDSG ASIRMQVYDR RIPLEYANKHT LLOKAAEFG
1501 AHY
11AA_SEQUENCE 1.0
ID AAG17202 standard; Protein; 584 AA.
XX
AC AAG17202;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18134.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152353.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.

PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 12-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 08-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.

PR 27-JAN-2000; 2000US-0178544.
 PR 27-JAN-2000; 2000US-0178545.
 PR 27-JAN-2000; 2000US-0178546.
 PR 27-JAN-2000; 2000US-0178547.
 PR 28-JAN-2000; 2000US-0178754.
 PR 28-JAN-2000; 2000US-0178755.
 PR 01-FEB-2000; 2000US-0179388.
 PR 01-FEB-2000; 2000US-0179389.
 PR 03-FEB-2000; 2000US-0180039.
 PR 03-FEB-2000; 2000US-0180139.
 PR 04-FEB-2000; 2000US-0180206.
 PR 04-FEB-2000; 2000US-0180207.
 PR 07-FEB-2000; 2000US-0180695.
 PR 07-FEB-2000; 2000US-0180696.
 PR 09-FEB-2000; 2000US-0181214.
 PR 09-FEB-2000; 2000US-0181228.
 PR 10-FEB-2000; 2000US-0181476.
 PR 10-FEB-2000; 2000US-0181551.
 PR 15-FEB-2000; 2000US-0182477.
 PR 15-FEB-2000; 2000US-0182478.
 PR 15-FEB-2000; 2000US-0182512.
 PR 15-FEB-2000; 2000US-0182516.
 PR 17-FEB-2000; 2000US-0183165.
 PR 17-FEB-2000; 2000US-0183166.

XX (CERE-) CERES INC.

PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
 PI Zheng L, Dumas J;

DR WPI: 2000-507395/46.
 DR N-PSDB; AAC37642.

PT New sequence determined DNA fragments (SDFs) from different plant
 PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
 PT protein coding sequences, untranslated regions, or as 3' termination
 PT sequences -

XX PS Claim 19; SEQ ID 18134; 344pp + CD-ROM; English.

XX The present sequence is a putative protein fragment from
 CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
 CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
 CC library which could then be sequenced, allowing the putative protein
 CC sequence(s) to be obtained. This sequence may be useful for protein
 CC identification and for aiding in the elucidation of signal transduction
 CC and metabolic pathways. Its coding sequence has a use in the control of
 CC gene expression as a promoter, coding sequence, 3'UTR or termination
 CC sequence, for controlling the behaviour of a gene within the chromosome,
 CC as a tool for use in genetic mapping, including a use in hybridisation
 CC assays, for recognition or isolation of similar DNA fragments, or for
 CC the identification of a particular organism.

XX Sequence 584 AA;

AA017202 Length: 584 May 30, 2002 09:25 Type: P Check: 7959

1 YFEIYVLLF LKIRONSSE KLFIVVEEF WSIYMGREG VSLADGRF
 51 HPTDEELVRY YLKRKICNRP FKFDIAISYTD VYKSEPMDLR DKSRLSRDL
 101 EWYFFSMLNK KYRNGSKTNR ATEMGYKTT GKOREILNGS KYVGMKTLV
 151 YHKGRAPGE RINWYHNEYR LVDODLDKGTG VHODAFVLCR IFOKSSGGPK
 201 NGDYGAPVY EEWEEDEDM TFVPDQEDLG SEDHVYVHMD DIQKSENFV
 251 VYDAIPPLN FIHGESSNNV EFNYSDSINY IQGTGNMDS GGYFEPAPES
 301 YEKDQKTIIR HRDGLQNEG IGCYQDKHS EIILOSSNIF GDTTSCYDNF
 351 PVESNYLIGE AFLDPSNLL ENDGLYLETN DLSSTQDGF DFEDYLTFFD

401 ETFPDQIMG NEDVFPDQEE LFQEVETKEL EKEETSRSKH VVEKEKDEA
 451 SCSKQVDADA TEFEPPDYKYP ILKKASHMLG AIPAPLANAS EFPKDAIR
 501 LHAQSSGSV HYTAGMITTS DSNMGWSYCK NENIDLITSL GLVQNTAPE
 551 KSGNSSAMAM LIPMCFWILL LSVSRKVSIL VSSR

ID IIA_SEQUENCE 1.0
 ID AAG17203 standard; Protein: 549 AA.

XX AAG17203;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 18135.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130571.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 14-JUN-1999; 99US-0138847.
 PR 16-JUN-1999; 99US-0139119.
 PR 99US-0139452.

PR 16-JUN-1999;	990S-0139453.
PR 17-JUN-1999;	990S-0139492.
PR 18-JUN-1999;	990S-0139454.
PR 18-JUN-1999;	990S-0139455.
PR 18-JUN-1999;	990S-0139456.
PR 18-JUN-1999;	990S-0139457.
PR 18-JUN-1999;	990S-0139458.
PR 18-JUN-1999;	990S-0139459.
PR 18-JUN-1999;	990S-0139460.
PR 18-JUN-1999;	990S-0139461.
PR 18-JUN-1999;	990S-0139462.
PR 18-JUN-1999;	990S-0139463.
PR 18-JUN-1999;	990S-0139750.
PR 21-JUN-1999;	990S-0139763.
PR 21-JUN-1999;	990S-0139817.
PR 22-JUN-1999;	990S-0139899.
PR 23-JUN-1999;	990S-0140353.
PR 23-JUN-1999;	990S-0140354.
PR 24-JUN-1999;	990S-0140695.
PR 28-JUN-1999;	990S-0140823.
PR 29-JUN-1999;	990S-0140991.
PR 30-JUN-1999;	990S-0141287.
PR 01-JUL-1999;	990S-0141842.
PR 01-JUL-1999;	990S-0142154.
PR 02-JUL-1999;	990S-0142055.
PR 06-JUL-1999;	990S-0142390.
PR 08-JUL-1999;	990S-0142803.
PR 09-JUL-1999;	990S-0142920.
PR 12-JUL-1999;	990S-0142977.
PR 13-JUL-1999;	990S-0143542.
PR 14-JUL-1999;	990S-0143624.
PR 15-JUL-1999;	990S-0144005.
PR 16-JUL-1999;	990S-0144085.
PR 16-JUL-1999;	990S-0144086.
PR 19-JUL-1999;	990S-0144325.
PR 19-JUL-1999;	990S-0144331.
PR 19-JUL-1999;	990S-0144332.
PR 19-JUL-1999;	990S-0144333.
PR 19-JUL-1999;	990S-0144334.
PR 19-JUL-1999;	990S-0144335.
PR 19-JUL-1999;	990S-0144352.
PR 20-JUL-1999;	990S-0144632.
PR 20-JUL-1999;	990S-0144884.
PR 21-JUL-1999;	990S-0144814.
PR 21-JUL-1999;	990S-0145086.
PR 21-JUL-1999;	990S-0145088.
PR 22-JUL-1999;	990S-0145085.
PR 22-JUL-1999;	990S-0145087.
PR 22-JUL-1999;	990S-0145089.
PR 22-JUL-1999;	990S-0145192.
PR 23-JUL-1999;	990S-0145145.
PR 23-JUL-1999;	990S-0145218.
PR 23-JUL-1999;	990S-0145224.
PR 26-JUL-1999;	990S-0145276.
PR 27-JUL-1999;	990S-0145913.
PR 27-JUL-1999;	990S-0145918.
PR 27-JUL-1999;	990S-0145919.
PR 28-JUL-1999;	990S-0145951.
PR 02-AUG-1999;	990S-0146386.
PR 02-AUG-1999;	990S-0146388.
PR 02-AUG-1999;	990S-0146389.
PR 03-AUG-1999;	990S-0147008.
PR 04-AUG-1999;	990S-0147204.
PR 04-AUG-1999;	990S-0147302.
PR 05-AUG-1999;	990S-0147192.
PR 05-AUG-1999;	990S-0147260.
PR 06-AUG-1999;	990S-0147303.
PR 06-AUG-1999;	990S-0147416.
PR 09-AUG-1999;	990S-0147493.
PR 09-AUG-1999;	990S-0147935.
PR 10-AUG-1999;	990S-0148171.
PR 11-AUG-1999;	990S-0148319.
PR 12-AUG-1999;	990S-0148341.
PR 13-AUG-1999;	990S-0148565.
PR 13-AUG-1999;	990S-0148684.
PR 16-AUG-1999;	990S-0149368.
PR 17-AUG-1999;	990S-0149175.
PR 18-AUG-1999;	990S-0149426.
PR 20-AUG-1999;	990S-0149722.
PR 20-AUG-1999;	990S-0149723.
PR 20-AUG-1999;	990S-0149929.
PR 23-AUG-1999;	990S-0149902.
PR 23-AUG-1999;	990S-0149930.
PR 25-AUG-1999;	990S-0150566.
PR 26-AUG-1999;	990S-0150884.
PR 27-AUG-1999;	990S-0151065.
PR 27-AUG-1999;	990S-0151080.
PR 30-AUG-1999;	990S-0151303.
PR 31-AUG-1999;	990S-0151438.
PR 01-SEP-1999;	990S-0151930.
PR 07-SEP-1999;	990S-0152363.
PR 10-SEP-1999;	990S-0153070.
PR 13-SEP-1999;	990S-0153758.
PR 15-SEP-1999;	990S-0154018.
PR 16-SEP-1999;	990S-0154039.
PR 20-SEP-1999;	990S-0154779.
PR 22-SEP-1999;	990S-0155139.
PR 23-SEP-1999;	990S-0155486.
PR 24-SEP-1999;	990S-0155659.
PR 28-SEP-1999;	990S-0156458.
PR 29-SEP-1999;	990S-0156596.
PR 04-OCT-1999;	990S-0157117.
PR 05-OCT-1999;	990S-0157753.
PR 06-OCT-1999;	990S-0157865.
PR 07-OCT-1999;	990S-0158029.
PR 08-OCT-1999;	990S-0158232.
PR 12-OCT-1999;	990S-0158369.
PR 13-OCT-1999;	990S-0159293.
PR 13-OCT-1999;	990S-0159294.
PR 13-OCT-1999;	990S-0159329.
PR 14-OCT-1999;	990S-0159330.
PR 14-OCT-1999;	990S-0159331.
PR 14-OCT-1999;	990S-0159637.
PR 14-OCT-1999;	990S-0159638.
PR 18-OCT-1999;	990S-0159584.
PR 21-OCT-1999;	990S-0160741.
PR 21-OCT-1999;	990S-0160767.
PR 21-OCT-1999;	990S-0160768.
PR 21-OCT-1999;	990S-0160770.
PR 21-OCT-1999;	990S-0160814.
PR 21-OCT-1999;	990S-0160815.
PR 22-OCT-1999;	990S-0160980.
PR 22-OCT-1999;	990S-0160981.
PR 22-OCT-1999;	990S-0160989.
PR 25-OCT-1999;	990S-0161404.
PR 25-OCT-1999;	990S-0161405.
PR 25-OCT-1999;	990S-0161406.
PR 26-OCT-1999;	990S-0161359.
PR 26-OCT-1999;	990S-0161360.
PR 26-OCT-1999;	990S-0161361.
PR 28-OCT-1999;	990S-0161920.
PR 28-OCT-1999;	990S-0161992.
PR 28-OCT-1999;	990S-0161993.
PR 29-OCT-1999;	990S-0162142.
PR 29-OCT-1999;	990S-0162143.
PR 29-OCT-1999;	990S-0162228.
PR 01-NOV-1999;	990S-0162891.
PR 01-NOV-1999;	990S-0162894.
PR 01-NOV-1999;	990S-0162895.
PR 02-NOV-1999;	990S-0163091.
PR 02-NOV-1999;	990S-0163092.
PR 02-NOV-1999;	990S-0163093.
PR 03-NOV-1999;	990S-0163248.
PR 03-NOV-1999;	990S-0163249.

PR 03-NOV-1999; 99US-0163381.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 08-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 09-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 10-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 12-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164570.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 16-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165861.
PR 16-NOV-1999; 99US-0165869.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 22-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 23-NOV-1999; 99US-0166750.
PR 24-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 30-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.

PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 15-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183185.
PR 17-FEB-2000; 2000US-0183186.

(CERE-) CERES INC.
XX
XX
PA Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
XX WPI; 2000-507395/46.
DR N-PDB; AAC37642.
XX
PT New sequence determined DNA fragments (SDRs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
XX
PS Claim 19; SEQ ID 18135; 344pp + CD-ROM; English.
XX
CC The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
SQ Sequence 549 AA;

AAC17203 Length: 549 May 30, 2002 09:25 Type: P Check: 1255 ..

1 MGRGSVTSLSA PGRFRPPTDE ELVRYLLKRR ICKPFFDA ISTVDYKSE
51 PMDLPDKSRL KSRDLEWYFF SMLNKKRYNG SKTNRATENG YWTKTGKRE
101 ILNGSKVYGM KKTLYYHKGR APRGERTMW MHEVRLVDD LKGTGHODA
151 FVLCRIFFOKS GSGPRKNGEY GAPFVEEME EEDMTVPVD QEDLGSEDHY
201 YVHMDIDQK SENFVYVDAL PIPLNFINGE SSNNVETNYS DSINTYQGTG
251 NYMDSGGYFE OPAESYEKQD KPIIRHRDGS LQNGIGICGV QDKHSIIDS
301 SDNIFGTDTS CYNDPVEESN YLIGEFALDP NSNLENDGL YETNDLSST
351 QODGFDEFDY LTFDETFDP SOLMGNEDEV FDOEELFOEV ETEKELEKEET
401 SRSKHYVEEK EKDEASCSKO VDADATEREP DYKYPILKKA SHMLGAIAPAP
451 LANASEPTPK DAIRILHAAQ SSGSVHYTAG MTIISDSNMG WSYGNENLID

501 LIISLGLVGG NTAPKSGNS SAMAMLIFMC FWWLLLSVSF KVSILVSSR
11A_SEQUENCE 1.0
ID ANGI7204 standard; Protein: 478 AA.
XX ANGI7204;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18136.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126284.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130891.
PR 30-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 04-MAY-1999; 990S-0132407.
PR 05-MAY-1999; 990S-0132484.
PR 06-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 11-MAY-1999; 990S-0132863.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.

PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142970.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147182.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147403.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.

PR 20-AUG-1999. 99US-0149723.
PR 20-AUG-1999. 99US-0149929.
PR 23-AUG-1999. 99US-0149902.
PR 23-AUG-1999. 99US-0149930.
PR 25-AUG-1999. 99US-0150566.
PR 26-AUG-1999. 99US-0150884.
PR 27-AUG-1999. 99US-0151065.
PR 27-AUG-1999. 99US-0151066.
PR 27-AUG-1999. 99US-0151080.
PR 30-AUG-1999. 99US-0151303.
PR 31-AUG-1999. 99US-0151438.
PR 01-SEP-1999. 99US-0151930.
PR 07-SEP-1999. 99US-0152363.
PR 10-SEP-1999. 99US-0153070.
PR 13-SEP-1999. 99US-0153758.
PR 15-SEP-1999. 99US-0154018.
PR 16-SEP-1999. 99US-0154039.
PR 20-SEP-1999. 99US-0154779.
PR 22-SEP-1999. 99US-0155139.
PR 23-SEP-1999. 99US-0155486.
PR 24-SEP-1999. 99US-0155659.
PR 28-SEP-1999. 99US-0156458.
PR 29-SEP-1999. 99US-0156596.
PR 04-OCT-1999. 99US-0157117.
PR 05-OCT-1999. 99US-0157753.
PR 06-OCT-1999. 99US-0157865.
PR 07-OCT-1999. 99US-0158029.
PR 08-OCT-1999. 99US-0158232.
PR 12-OCT-1999. 99US-0158369.
PR 13-OCT-1999. 99US-0159293.
PR 13-OCT-1999. 99US-0159294.
PR 13-OCT-1999. 99US-0159295.
PR 14-OCT-1999. 99US-0159329.
PR 14-OCT-1999. 99US-0159330.
PR 14-OCT-1999. 99US-0159331.
PR 14-OCT-1999. 99US-0159337.
PR 14-OCT-1999. 99US-0159637.
PR 14-OCT-1999. 99US-0159638.
PR 18-OCT-1999. 99US-0159584.
PR 21-OCT-1999. 99US-0160741.
PR 21-OCT-1999. 99US-0160767.
PR 21-OCT-1999. 99US-0160768.
PR 21-OCT-1999. 99US-0160770.
PR 21-OCT-1999. 99US-0160814.
PR 21-OCT-1999. 99US-0160815.
PR 22-OCT-1999. 99US-0160880.
PR 22-OCT-1999. 99US-0160981.
PR 22-OCT-1999. 99US-0160989.
PR 25-OCT-1999. 99US-0161404.
PR 25-OCT-1999. 99US-0161405.
PR 25-OCT-1999. 99US-0161406.
PR 26-OCT-1999. 99US-0161359.
PR 26-OCT-1999. 99US-0161360.
PR 26-OCT-1999. 99US-0161361.
PR 28-OCT-1999. 99US-0161920.
PR 28-OCT-1999. 99US-0161992.
PR 28-OCT-1999. 99US-0161993.
PR 29-OCT-1999. 99US-0162142.
PR 29-OCT-1999. 99US-0162143.
PR 29-OCT-1999. 99US-0162228.
PR 01-NOV-1999. 99US-0162891.
PR 01-NOV-1999. 99US-0162894.
PR 01-NOV-1999. 99US-0162895.
PR 02-NOV-1999. 99US-0163091.
PR 02-NOV-1999. 99US-0163092.
PR 02-NOV-1999. 99US-0163093.
PR 03-NOV-1999. 99US-0163248.
PR 03-NOV-1999. 99US-0163249.
PR 03-NOV-1999. 99US-0163281.
PR 04-NOV-1999. 99US-0163379.
PR 04-NOV-1999. 99US-0163380.
PR 04-NOV-1999. 99US-0163381.
PR 08-NOV-1999. 99US-0164146.
PR 08-NOV-1999. 99US-0164150.

PR 08-NOV-1999. 99US-0164151.
PR 09-NOV-1999. 99US-0164259.
PR 09-NOV-1999. 99US-0164260.
PR 10-NOV-1999. 99US-0164317.
PR 10-NOV-1999. 99US-0164318.
PR 10-NOV-1999. 99US-0164319.
PR 10-NOV-1999. 99US-0164321.
PR 10-NOV-1999. 99US-0164544.
PR 10-NOV-1999. 99US-0164545.
PR 12-NOV-1999. 99US-0164870.
PR 12-NOV-1999. 99US-0164871.
PR 12-NOV-1999. 99US-0164859.
PR 12-NOV-1999. 99US-0164960.
PR 12-NOV-1999. 99US-0164961.
PR 12-NOV-1999. 99US-0164962.
PR 15-NOV-1999. 99US-0164926.
PR 15-NOV-1999. 99US-0164927.
PR 15-NOV-1999. 99US-0164929.
PR 16-NOV-1999. 99US-0165661.
PR 16-NOV-1999. 99US-0165669.
PR 16-NOV-1999. 99US-0165671.
PR 17-NOV-1999. 99US-0165911.
PR 17-NOV-1999. 99US-0165918.
PR 17-NOV-1999. 99US-0165919.
PR 18-NOV-1999. 99US-0166157.
PR 18-NOV-1999. 99US-0166158.
PR 18-NOV-1999. 99US-0166173.
PR 19-NOV-1999. 99US-0166411.
PR 19-NOV-1999. 99US-0166412.
PR 19-NOV-1999. 99US-0166419.
PR 22-NOV-1999. 99US-0166733.
PR 22-NOV-1999. 99US-0166735.
PR 23-NOV-1999. 99US-0167362.
PR 24-NOV-1999. 99US-0167233.
PR 24-NOV-1999. 99US-0167234.
PR 24-NOV-1999. 99US-0167335.
PR 30-NOV-1999. 99US-0167382.
PR 30-NOV-1999. 99US-0167902.
PR 30-NOV-1999. 99US-0167904.
PR 01-DEC-1999. 99US-0168231.
PR 01-DEC-1999. 99US-0168232.
PR 01-DEC-1999. 99US-0168233.
PR 02-DEC-1999. 99US-0168546.
PR 02-DEC-1999. 99US-0168548.
PR 03-DEC-1999. 99US-0168549.
PR 03-DEC-1999. 99US-0168673.
PR 03-DEC-1999. 99US-0168674.
PR 03-DEC-1999. 99US-0168675.
PR 07-DEC-1999. 99US-0169278.
PR 07-DEC-1999. 99US-0169298.
PR 07-DEC-1999. 99US-0169302.
PR 08-DEC-1999. 99US-0169691.
PR 08-DEC-1999. 99US-0169692.
PR 16-DEC-1999. 99US-0171098.
PR 16-DEC-1999. 99US-0171107.
PR 16-DEC-1999. 99US-0171114.
PR 19-JAN-2000. 2000US-0176866.
PR 19-JAN-2000. 2000US-0176867.
PR 19-JAN-2000. 2000US-0176910.
PR 26-JAN-2000. 2000US-0178470.
PR 27-JAN-2000. 2000US-0178476.
PR 27-JAN-2000. 2000US-0178477.
PR 27-JAN-2000. 2000US-0178544.
PR 27-JAN-2000. 2000US-0178545.
PR 27-JAN-2000. 2000US-0178546.
PR 27-JAN-2000. 2000US-0178547.
PR 28-JAN-2000. 2000US-0178754.
PR 28-JAN-2000. 2000US-0178755.
PR 01-FEB-2000. 2000US-0179388.
PR 01-FEB-2000. 2000US-0179395.
PR 03-FEB-2000. 2000US-0180039.
PR 03-FEB-2000. 2000US-0180139.

```

PR      04-FEB-2000; 2000US-0180206.
PR      04-FEB-2000; 2000US-0180207.
PR      07-FEB-2000; 2000US-0180695.
PR      07-FEB-2000; 2000US-0180696.
PR      09-FEB-2000; 2000US-0181214.
PR      09-FEB-2000; 2000US-0181228.
PR      10-FEB-2000; 2000US-0181476.
PR      10-FEB-2000; 2000US-0181551.
PR      15-FEB-2000; 2000US-0182477.
PR      15-FEB-2000; 2000US-0182478.
PR      15-FEB-2000; 2000US-0182512.
PR      15-FEB-2000; 2000US-0182516.
PR      17-FEB-2000; 2000US-0183165.
PR      17-FEB-2000; 2000US-0183166.
XX
XX      (CERE-) CERES INC.
PA
XX      Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI      Zheng L, Dumas J;
XX
XX      WPI: 2000-507395/46.
DR      N-PSDB: AAC37642.
XX
XX
XX      New sequence determined DNA fragments (SDS) from different plant
PT      species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT      protein coding sequences, untranslated regions, or as 3' termination
PT      sequences -
XX
PS      Claim 19; SEQ ID 18136; 344pp + CD-ROM; English.
XX
XX      The present sequence is a putative protein fragment from
CC      Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC      RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC      library which could then be sequenced, allowing the putative protein
CC      sequence(s) to be obtained. This sequence may be useful for protein
CC      identification and for aiding in the elucidation of signal transduction
CC      and metabolic pathways. Its coding sequence has a use in the control of
CC      gene expression as a promoter, coding sequence, 3'UTR or termination
CC      sequence, for controlling the behaviour of a gene within the chromosome
CC      as a tool for use in genetic mapping, including a use in hybridisation
CC      assays, for recognition or isolation of similar DNA fragments, or for
CC      the identification of a particular organism.
XX
XX      Sequence    478 AA:
SO
AAAG17204 Length: 478 May 30, 2002 09:25 Type: P Check: 530 ..
1 MLNKKRYRGS KTNRATEMGY WKTKGRDREI LNGSKVVGAK KTLVYHKGRA
51 PRGETNTMYM HEYRLVDQDL DKTGVQDAF VLCRIFRKSG SGPRNGQYG
101 APVEEEMEE EDDMTFVPDQ EDLGSBDHY VHMDIDOKS ENFYVYDAIP
151 IPLNFHGES SNNETNYSD SINYIQOTGN YMSGGYFEQ PAESYEKDK
201 PIIRROGSL QNEGIGCVQ DKHSLEQSS DNIFETDSC YNDPPVASNY
251 LIGEFALDPN SNLLENDGLY LETNDLSSTQ ODGFDEEDYL TFEDETDFPS
301 QLMGNDEVFF DOEELFOEVE TKLEKEETS RSKHVEEKE KDEASCSROY
351 DADATEFPED YKYPLTKKAS HMLGAIPARL ANASEFTPKD AAIRLHAQS
401 SGSVAIVTGM ITISDSNMGW STCKNNLNL ILISGLVQGN TAPKSGNSS
451 AWAMLIENCF WVLLSVSFK VSILVSSR
IIAA_SEQUENCE 1.0
ID   AAG21029 standard; Protein: 504 AA.
AC   AAG21029;
XX
XX      17-OCT-2000 (first entry)
DT

```

XX	Arabidopsis thaliana protein fragment SEQ ID NO: 23437.
DE	
XX	Protein identification; signal transduction pathway; metabolic pathway;
KM	hybridisation assay; genetic mapping; gene expression control; promoter;
KM	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PM	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PE	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	23-APR-1999; 99US-0130891.
PR	28-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	30-APR-1999; 99US-0132487.
PR	04-MAY-1999; 99US-0132484.
PR	05-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	06-MAY-1999; 99US-0132487.
PR	07-MAY-1999; 99US-0132863.
PR	11-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.
PR	14-MAY-1999; 99US-0134370.
PR	14-MAY-1999; 99US-0134768.
PR	18-MAY-1999; 99US-0134941.
PR	19-MAY-1999; 99US-0135124.
PR	20-MAY-1999; 99US-0135353.
PR	21-MAY-1999; 99US-0135629.
PR	24-MAY-1999; 99US-0136021.
PR	25-MAY-1999; 99US-0136392.
PR	27-MAY-1999; 99US-0136782.
PR	28-MAY-1999; 99US-0137222.
PR	01-JUN-1999; 99US-0137528.
PR	03-JUN-1999; 99US-0137502.
PR	04-JUN-1999; 99US-0137724.
PR	07-JUN-1999; 99US-0138094.
PR	08-JUN-1999; 99US-0138540.
PR	10-JUN-1999; 99US-0138847.
PR	10-JUN-1999; 99US-0139119.
PR	14-JUN-1999; 99US-0139452.
PR	16-JUN-1999; 99US-0139453.
PR	16-JUN-1999; 99US-0139459.
PR	17-JUN-1999; 99US-0139460.
PR	18-JUN-1999; 99US-0139461.
PR	18-JUN-1999; 99US-0139456.
PR	18-JUN-1999; 99US-0139457.
PR	18-JUN-1999; 99US-0139458.
PR	18-JUN-1999; 99US-0139459.
PR	18-JUN-1999; 99US-0139462.
PR	18-JUN-1999; 99US-0139463.
PR	18-JUN-1999; 99US-0139750.
PR	18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.

```

PR 10-NOV-1999; 9905-0164545.
PR 10-NOV-1999; 9905-0164548.
PR 12-NOV-1999; 9905-0164870.
PR 12-NOV-1999; 9905-0164871.
PR 12-NOV-1999; 9905-0164859.
PR 12-NOV-1999; 9905-0164960.
PR 12-NOV-1999; 9905-0164961.
PR 12-NOV-1999; 9905-0164962.
PR 15-NOV-1999; 9905-0164926.
PR 15-NOV-1999; 9905-0164927.
PR 15-NOV-1999; 9905-0164929.
PR 16-NOV-1999; 9905-0165661.
PR 16-NOV-1999; 9905-0165669.
PR 16-NOV-1999; 9905-0165671.
PR 17-NOV-1999; 9905-0165911.
PR 17-NOV-1999; 9905-0165918.
PR 17-NOV-1999; 9905-0165919.
PR 18-NOV-1999; 9905-0166157.
PR 18-NOV-1999; 9905-0166158.
PR 18-NOV-1999; 9905-0166173.
PR 19-NOV-1999; 9905-0166411.
PR 19-NOV-1999; 9905-0166412.
PR 19-NOV-1999; 9905-0166419.
PR 22-NOV-1999; 9905-0166733.
PR 22-NOV-1999; 9905-0166750.
PR 23-NOV-1999; 9905-0167262.
PR 24-NOV-1999; 9905-0167233.
PR 24-NOV-1999; 9905-0167234.
PR 24-NOV-1999; 9905-0167235.
PR 24-NOV-1999; 9905-0167382.
PR 30-NOV-1999; 9905-0167902.
PR 30-NOV-1999; 9905-0167904.
PR 30-NOV-1999; 9905-0167908.
PR 01-DEC-1999; 9905-0168231.
PR 01-DEC-1999; 9905-0168232.
PR 01-DEC-1999; 9905-0168233.
PR 02-DEC-1999; 9905-0168546.
PR 02-DEC-1999; 9905-0168548.
PR 02-DEC-1999; 9905-0168549.
PR 03-DEC-1999; 9905-0168673.
PR 03-DEC-1999; 9905-0168674.
PR 03-DEC-1999; 9905-0168675.
PR 07-DEC-1999; 9905-0169278.
PR 07-DEC-1999; 9905-0169298.
PR 08-DEC-1999; 9905-0169302.
PR 08-DEC-1999; 9905-0169691.
PR 16-DEC-1999; 9905-0169692.
PR 16-DEC-1999; 9905-0171098.
PR 16-DEC-1999; 9905-0171107.
PR 16-DEC-1999; 9905-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.

```

```

PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.
XX
XX (CERE-) CERES INC.
XX
XX Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
XX WPI: 2000-507395/46.
DR N-PSDB: AAC39118.
XX
XX New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
XX Claim 19; SEQ ID 23437; 344pp + CD-ROM; English.
PS
XX The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3' UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
XX Sequence 504 AA:
SQ
AGG21029 Length: 504 May 30, 2002 09:25 Type: P Check: 5419 ..
1 MVTGGEFHV RRRDQNLNLSLQSFSLFW SMFSPLESS LSONPKISFP
51 LVFSEGLLC DGRVNTYST PAPHNSDAD PLNPTREBS HFSIALSKF
101 LEOFFIDTSF GSRSAVLAVT PLVPSHRISL MARDSLGSFV VPFPVSTST
151 HDELHCQGVF FPGASPTIT SDLYDGCALP SLSTVTTLG FCGGGLNPID
201 GPVATIHAFG GLGISQFGGP KTLNLIWPTS RFALLRVIST LSLNHCFIT
251 DGRGLHSIVG TLVLEPCNH NFRNLHRYD WLPTDVSLV LKSLGLVLISN
301 HRLCDAAHAF VFIODHIMSL NLPTIMLTHP VPCRLDDAS QPSNMCWIKL
351 LRPALSTFS SPNNPVYLCH LIRLSVLMT CNLCCRLHVV YSSKEIDSLK
401 PYSSKEISSL KLEPGEKSHT AASLEHFNK SEVASIARAY YDKHLINDFS
451 KLVFKLESIQ VPIDPLSLAT FSKLCILIEDS WMLYLIQSN ALIILCINPP
501 LIRS
11AA_SEQUENCE 1.0
ID AAG26394 standard; Protein: 329 AA.
XX
XX AAG26394;
AC
XX
XX 17-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 30833.
DE
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hydriidation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

```

```
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000: 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 04-MAY-1999; 99US-0132048.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 19-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0134941.
XX 21-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0138119.
XX 16-JUN-1999; 99US-0138452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139452.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 21-JUN-1999; 99US-0139763.
XX 22-JUN-1999; 99US-0139817.
XX 23-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
PR
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
```

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158869.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.

PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164966.
PR 15-NOV-1999; 99US-0164967.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167235.
PR 30-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169681.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 19-JAN-2000; 2000US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.

XX (CERE-) CERES INC.
XX Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX MPI: 2000-507395/46.
DR N-PSDB; AAC41149.
XX
XX New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
PS Claim 19; SEQ ID 30833; 344bp + CD-ROM; English.
XX
XX The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
SQ Sequence 329 AA:
AAG26394 Length: 329 May 30, 2002 09:25 Type: P Check: 3190 ..
1 MWOSQSLSLT TICGSVKYSS LLMRLNSVK ASSLIDRCV SCQFLRKSPS
51 FFSHMSLNQ RNLLEVEARW PFOGGGGLD PSSESESAN EDILIEFFOL
101 DLATROYAM NLEQYDIAQ LREKLEVEE ESIRLOEGRK GSSAKSEAD
151 KGISIRRLA DLQNAIDSED YGLAAKLDE ISKLEAESLA VSAKALAFKK
201 AEYAFRLGOK LRHKTFGYRA VVCGMDPICC ESSSMMEAAE VEKLPRGSNQ
251 PFYQVLVDVR THPDLVAVV AEDNLFSPKK PDKERFDHPY IEFLYGADIT
301 AGDFIPVKQL REKYNPRHE VPFDSDQED
11AA_SEQUENCE 1.0
ID AAG26395 standard; Protein; 307 AA.
XX
AC AAG26395;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 30834.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
OS EP1033405-A2.
XX
PN 06-SEP-2000.
XX
PD 25-FEB-2000; 2000EP-0301439.
XX
PF 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 9905-01443325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144332.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 22-JUL-1999; 9905-0145195.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145219.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145226.
PR 27-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147483.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148568.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154079.
PR 20-SEP-1999; 9905-0154739.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.

PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.
PR 29-OCT-1999; 9905-0162143.
PR 29-OCT-1999; 9905-0162228.
PR 01-NOV-1999; 9905-0162891.
PR 01-NOV-1999; 9905-0162894.
PR 01-NOV-1999; 9905-0162895.
PR 02-NOV-1999; 9905-0163091.
PR 02-NOV-1999; 9905-0163093.
PR 02-NOV-1999; 9905-0163094.
PR 03-NOV-1999; 9905-0163248.
PR 03-NOV-1999; 9905-0163249.
PR 03-NOV-1999; 9905-0163281.
PR 04-NOV-1999; 9905-0163379.
PR 04-NOV-1999; 9905-0163380.
PR 04-NOV-1999; 9905-0163381.
PR 08-NOV-1999; 9905-0164146.
PR 08-NOV-1999; 9905-0164150.
PR 08-NOV-1999; 9905-0164151.
PR 09-NOV-1999; 9905-0164259.
PR 09-NOV-1999; 9905-0164260.
PR 10-NOV-1999; 9905-0164317.
PR 10-NOV-1999; 9905-0164318.
PR 10-NOV-1999; 9905-0164319.
PR 10-NOV-1999; 9905-0164321.
PR 10-NOV-1999; 9905-0164344.
PR 10-NOV-1999; 9905-0164545.
PR 10-NOV-1999; 9905-0164548.
PR 12-NOV-1999; 9905-0164870.
PR 12-NOV-1999; 9905-0164871.
PR 12-NOV-1999; 9905-0164959.
PR 12-NOV-1999; 9905-0164960.
PR 12-NOV-1999; 9905-0164961.
PR 12-NOV-1999; 9905-0164962.
PR 15-NOV-1999; 9905-0164926.
PR 15-NOV-1999; 9905-0164927.
PR 15-NOV-1999; 9905-0164929.
PR 16-NOV-1999; 9905-0165661.
PR 16-NOV-1999; 9905-0165669.
PR 16-NOV-1999; 9905-0165671.
PR 17-NOV-1999; 9905-0165911.
PR 17-NOV-1999; 9905-0165918.
PR 17-NOV-1999; 9905-0165919.
PR 18-NOV-1999; 9905-0166157.
PR 18-NOV-1999; 9905-0166158.
PR 18-NOV-1999; 9905-0166159.
PR 18-NOV-1999; 9905-0166173.

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139751.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147182.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149398.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149446.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156556.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

XX Sequence 623 AA;
SQ
AG29842 Length: 623 May 30, 2002 09:25 Type: P Check: 7775 ..
1 DFRKSVNGDL ROPHESNMPR CATERDNETN FKSRYSHLL LCSLXPSIL
51 YCREEDDVE YKERTQKEK RRMSTLDSI DALLFSLRA FTSPPAFVQ
101 IQGCTICLL ALGMLAEVY RNRVARIKN SIKAGNSLAF LYQDINLEH
151 SROYKLPRVS VVMPKGFGE HNLHNRQOI TSLYGGPIEF LEVESTEDP
201 AYHVSRLLS MYODHVEAKV VVAGLSTTCS OKIHNLIGV EKMKDKRYV
251 LFLEDDVRLH PGTIGALTE MEKNPEVPCS MGATGGRTF FLWGGGMMH
301 ADDFRQDRYG VVSGLDGCV SDDMTLASLA GAHKRLITSP PVAVFPPLA
351 SDSLFGRYWN YLRKQTFVLE SYISKVNIM NKALFAVHCY LSMGFVAPV
401 MAIHHSAL RIYIKGYHOL EDTTSASGL SVFLMLHIDR RMREYGMML
451 VTPLICTFI ELISMNMLR REVOLCNMLS PEAPRLIAT YNMGLPPIIG
501 GHISPMYRGF NKPRTGVEWN VEDYPIPNLY PHTIFONIKS ALENKHNVE
551 TPLLQTYVSS SIMLSTSLFD RGLIKRPSFT SSAASMPPET MYKGLNKTS
601 LYLFRALION IHISNIAMSA RMC
11AA_SEQUENCE 1.0
ID AG29843 standard; Protein: 550 AA.
XX
AC AG29843;
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35575.
XX
KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127452.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141267.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145293.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 04-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 08-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 09-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164670.
PR 12-NOV-1999; 99US-0164671.
PR 12-NOV-1999; 99US-0164699.
PR 12-NOV-1999; 99US-0164860.
PR 12-NOV-1999; 99US-0164861.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166159.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166735.
PR 22-NOV-1999; 99US-0166736.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167335.
PR 30-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144816.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161406.


```

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164344.
PR 10-NOV-1999; 99US-0164345.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.

```

```

PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0178666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.

```

(CERE-) CERES INC.

PA Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;

WPI: 2000-507395/46.
N-PSDB; AAC42442.

DR New sequence determined DNA fragments (SDPs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
sequences -

XX Claim 19; SEQ ID 35576; 344pp + CD-ROM; English.

XX The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3' UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.

SQ Sequence 461 AA;

AAG29644 Length: 461 May 30, 2002 09:25 Type: P Check: 4529 ..

1 MPLKGFGEHN LHMNRQITS LYGGPLEPLF VVESTEDPAY HANSRLISMV

51 QDHVEAKVYV AGLSTCSQK IHNQLIGVER MKKDKRYVLF LDDVRLHPG

101 TIGALLTTEME KNPVEPCSMG FATGRTFFL WGGCMMHAD DFRQDRYGVV
151 SGLRDGYSY DMTLASLGA HKRLITSPV AVPPHPLASD LSPGRYWNYL
201 RKQTFVLESY ISKVMIMNK ALFAVHCYLS WGFAPYVMA IHIITSLARI
251 YIKGYHOLED TTSAGGSLV FLMLHIDRM RFIYGMMLVI TLAICTFIEL
301 LSMWNLTTRRE VOLCNMLSPE APRLSLATYN WGLPFIIGH ISPMYRGFNK
351 PRTEGVFWNE DYPILPLYPH TIFONIKSAL ENKHFVETP LILQTFVSSSI
401 WLSTSLFDRC ILIKPSFTSS AASNPETMY KGLNGKTSLS LFRALQNIH
451 ISNIAMSARW C
11AA SEQUENCE 1.0
ID AAG36897 standard; Protein; 333 AA.
AC AAG36897;
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 45282.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 45282.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144634.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154039.
PR 16-SEP-1999;	99US-0154018.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
PR 21-OCT-1999;	99US-0160815.
PR 22-OCT-1999;	99US-0160960.
PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160989.
PR 25-OCT-1999;	99US-0161404.
PR 25-OCT-1999;	99US-0161405.
PR 25-OCT-1999;	99US-0161406.
PR 26-OCT-1999;	99US-0161359.
PR 26-OCT-1999;	99US-0161360.
PR 26-OCT-1999;	99US-0161361.
PR 28-OCT-1999;	99US-0161920.
PR 28-OCT-1999;	99US-0161992.
PR 28-OCT-1999;	99US-0161993.
PR 29-OCT-1999;	99US-0162142.
PR 29-OCT-1999;	99US-0162143.

PR 29-OCT-1999;	99US-0162228.
PR 01-NOV-1999;	99US-0162891.
PR 01-NOV-1999;	99US-0162894.
PR 01-NOV-1999;	99US-0162895.
PR 02-NOV-1999;	99US-0163091.
PR 02-NOV-1999;	99US-0163092.
PR 02-NOV-1999;	99US-0163093.
PR 03-NOV-1999;	99US-0163248.
PR 03-NOV-1999;	99US-0163249.
PR 04-NOV-1999;	99US-0163281.
PR 04-NOV-1999;	99US-0163379.
PR 04-NOV-1999;	99US-0163380.
PR 08-NOV-1999;	99US-0163381.
PR 08-NOV-1999;	99US-0164146.
PR 08-NOV-1999;	99US-0164150.
PR 08-NOV-1999;	99US-0164151.
PR 09-NOV-1999;	99US-0164259.
PR 09-NOV-1999;	99US-0164260.
PR 10-NOV-1999;	99US-0164317.
PR 10-NOV-1999;	99US-0164318.
PR 10-NOV-1999;	99US-0164319.
PR 10-NOV-1999;	99US-0164321.
PR 10-NOV-1999;	99US-0164344.
PR 10-NOV-1999;	99US-0164545.
PR 10-NOV-1999;	99US-0164548.
PR 12-NOV-1999;	99US-0164870.
PR 12-NOV-1999;	99US-0164871.
PR 12-NOV-1999;	99US-0164959.
PR 12-NOV-1999;	99US-0164960.
PR 12-NOV-1999;	99US-0164961.
PR 12-NOV-1999;	99US-0164962.
PR 15-NOV-1999;	99US-0164927.
PR 15-NOV-1999;	99US-0164929.
PR 16-NOV-1999;	99US-0165861.
PR 16-NOV-1999;	99US-0165669.
PR 16-NOV-1999;	99US-0165671.
PR 17-NOV-1999;	99US-0165911.
PR 17-NOV-1999;	99US-0165918.
PR 17-NOV-1999;	99US-0165919.
PR 18-NOV-1999;	99US-0166157.
PR 18-NOV-1999;	99US-0166158.
PR 18-NOV-1999;	99US-0166173.
PR 19-NOV-1999;	99US-0166411.
PR 19-NOV-1999;	99US-0166412.
PR 19-NOV-1999;	99US-0166419.
PR 22-NOV-1999;	99US-0166733.
PR 22-NOV-1999;	99US-0166750.
PR 23-NOV-1999;	99US-0167362.
PR 24-NOV-1999;	99US-0167233.
PR 24-NOV-1999;	99US-0167234.
PR 24-NOV-1999;	99US-0167235.
PR 24-NOV-1999;	99US-0167382.
PR 30-NOV-1999;	99US-0167902.
PR 30-NOV-1999;	99US-0167904.
PR 30-NOV-1999;	99US-0167908.
PR 01-DEC-1999;	99US-0168231.
PR 01-DEC-1999;	99US-0168232.
PR 01-DEC-1999;	99US-0168233.
PR 02-DEC-1999;	99US-0168546.
PR 02-DEC-1999;	99US-0168548.
PR 03-DEC-1999;	99US-0168549.
PR 03-DEC-1999;	99US-0168673.
PR 03-DEC-1999;	99US-0168674.
PR 03-DEC-1999;	99US-0168675.
PR 07-DEC-1999;	99US-0169278.
PR 07-DEC-1999;	99US-0169298.
PR 07-DEC-1999;	99US-0169302.
PR 08-DEC-1999;	99US-0169691.
PR 08-DEC-1999;	99US-0169692.
PR 16-DEC-1999;	99US-0171098.
PR 16-DEC-1999;	99US-0171107.
PR 16-DEC-1999;	99US-0171114.

PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180695.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.
PA (CERE-) CERES INC.
XX
XX
PI Alexandrov N, Brower V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
XX
DR MPI; 2000-507395/46.
DR N-PSDB; AAC45095.
XX
XX
PT New sequence determined DNA fragments (SPFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
XX
PS Claim 19; SEQ ID 45282; 344pp + CD-ROM, English.
XX
XX
CC The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for adding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
XX
SQ Sequence 333 AA;
AAG36897 Length: 333 May 30, 2002 09:25 Type: P Check: 2209 ..

301 SSPDEEDHWH TLDYKRCIQ GISNLLDQKN MKT
IDAA_SEQUENCE 1.0
ID AAG36898 standard; Protein: 318 AA.
XX
XX AAG36898;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 45283.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
PN EP1033405-A2.
XX
XX
XX 06-SEP-2000.
PD
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
PE
XX
XX 25-FEB-1999; 990US-0121825.
PR 03-MAR-1999; 990US-0123180.
PR 09-MAR-1999; 990US-0123548.
PR 23-MAR-1999; 990US-0125788.
PR 25-MAR-1999; 990US-0126264.
PR 29-MAR-1999; 990US-0126785.
PR 01-APR-1999; 990US-0127462.
PR 06-APR-1999; 990US-0128234.
PR 08-APR-1999; 990US-0128714.
PR 16-APR-1999; 990US-0129845.
PR 19-APR-1999; 990US-0130077.
PR 21-APR-1999; 990US-0130449.
PR 23-APR-1999; 990US-0130510.
PR 23-APR-1999; 990US-0130891.
PR 28-APR-1999; 990US-0131449.
PR 30-APR-1999; 990US-0132048.
PR 30-APR-1999; 990US-0132407.
PR 04-MAY-1999; 990US-0132484.
PR 05-MAY-1999; 990US-0132485.
PR 06-MAY-1999; 990US-0132486.
PR 06-MAY-1999; 990US-0132487.
PR 07-MAY-1999; 990US-0132863.
PR 11-MAY-1999; 990US-0134256.
PR 14-MAY-1999; 990US-0134218.
PR 14-MAY-1999; 990US-0134219.
PR 14-MAY-1999; 990US-0134221.
PR 14-MAY-1999; 990US-0134370.
PR 18-MAY-1999; 990US-0134768.
PR 19-MAY-1999; 990US-0134941.
PR 20-MAY-1999; 990US-0135124.
PR 21-MAY-1999; 990US-0135524.
PR 24-MAY-1999; 990US-0135629.
PR 25-MAY-1999; 990US-0136021.
PR 27-MAY-1999; 990US-0136392.
PR 28-MAY-1999; 990US-0136782.
PR 01-JUN-1999; 990US-0137222.
PR 03-JUN-1999; 990US-0137528.
PR 04-JUN-1999; 990US-0137502.
PR 07-JUN-1999; 990US-0137724.
PR 08-JUN-1999; 990US-0138094.
PR 10-JUN-1999; 990US-0138540.
PR 10-JUN-1999; 990US-0138847.
PR 14-JUN-1999; 990US-0139119.
PR 16-JUN-1999; 990US-0139452.
PR 16-JUN-1999; 990US-0139453.
PR 17-JUN-1999; 990US-0139492.
PR 18-JUN-1999; 990US-0139454.
PR 18-JUN-1999; 990US-0139455.
PR 18-JUN-1999; 990US-0139456.

PR 18-JUN-1999;	9905-0139457.
PR 18-JUN-1999;	9905-0139458.
PR 18-JUN-1999;	9905-0139459.
PR 18-JUN-1999;	9905-0139460.
PR 18-JUN-1999;	9905-0139461.
PR 18-JUN-1999;	9905-0139462.
PR 18-JUN-1999;	9905-0139463.
PR 18-JUN-1999;	9905-0139750.
PR 21-JUN-1999;	9905-0139763.
PR 22-JUN-1999;	9905-0139817.
PR 22-JUN-1999;	9905-0139889.
PR 23-JUN-1999;	9905-0140353.
PR 24-JUN-1999;	9905-0140354.
PR 28-JUN-1999;	9905-0140695.
PR 29-JUN-1999;	9905-0140823.
PR 30-JUN-1999;	9905-0140991.
PR 01-JUL-1999;	9905-0141287.
PR 01-JUL-1999;	9905-0141842.
PR 02-JUL-1999;	9905-0142154.
PR 06-JUL-1999;	9905-0142055.
PR 08-JUL-1999;	9905-0142390.
PR 09-JUL-1999;	9905-0142803.
PR 12-JUL-1999;	9905-0142920.
PR 13-JUL-1999;	9905-0142977.
PR 14-JUL-1999;	9905-0143542.
PR 15-JUL-1999;	9905-0143624.
PR 16-JUL-1999;	9905-0144005.
PR 16-JUL-1999;	9905-0144086.
PR 16-JUL-1999;	9905-0144335.
PR 19-JUL-1999;	9905-0144331.
PR 19-JUL-1999;	9905-0144332.
PR 19-JUL-1999;	9905-0144333.
PR 19-JUL-1999;	9905-0144334.
PR 20-JUL-1999;	9905-0144335.
PR 20-JUL-1999;	9905-0144632.
PR 21-JUL-1999;	9905-0144884.
PR 21-JUL-1999;	9905-0145086.
PR 21-JUL-1999;	9905-0145088.
PR 22-JUL-1999;	9905-0145085.
PR 22-JUL-1999;	9905-0145087.
PR 22-JUL-1999;	9905-0145089.
PR 23-JUL-1999;	9905-0145192.
PR 23-JUL-1999;	9905-0145218.
PR 26-JUL-1999;	9905-0145224.
PR 27-JUL-1999;	9905-0145276.
PR 27-JUL-1999;	9905-0145913.
PR 28-JUL-1999;	9905-0145918.
PR 28-JUL-1999;	9905-0145951.
PR 02-AUG-1999;	9905-0146386.
PR 02-AUG-1999;	9905-0146388.
PR 03-AUG-1999;	9905-0146389.
PR 04-AUG-1999;	9905-0147038.
PR 04-AUG-1999;	9905-0147204.
PR 05-AUG-1999;	9905-0147302.
PR 05-AUG-1999;	9905-0147192.
PR 06-AUG-1999;	9905-0147260.
PR 06-AUG-1999;	9905-0147303.
PR 09-AUG-1999;	9905-0147416.
PR 09-AUG-1999;	9905-0147493.
PR 10-AUG-1999;	9905-0147935.
PR 11-AUG-1999;	9905-0148171.
PR 12-AUG-1999;	9905-0148319.
PR 13-AUG-1999;	9905-0148341.
PR 13-AUG-1999;	9905-0148565.
PR 15-AUG-1999;	9905-0148684.
PR 17-AUG-1999;	9905-0149358.
PR 18-AUG-1999;	9905-0149175.
PR 18-AUG-1999;	9905-0149426.

PR 20-AUG-1999;	9905-0149722.
PR 20-AUG-1999;	9905-0149723.
PR 20-AUG-1999;	9905-0149929.
PR 23-AUG-1999;	9905-0149902.
PR 23-AUG-1999;	9905-0149930.
PR 25-AUG-1999;	9905-0150566.
PR 26-AUG-1999;	9905-0150884.
PR 27-AUG-1999;	9905-0151065.
PR 27-AUG-1999;	9905-0151066.
PR 30-AUG-1999;	9905-0151080.
PR 31-AUG-1999;	9905-0151303.
PR 01-SEP-1999;	9905-0151438.
PR 07-SEP-1999;	9905-0151930.
PR 10-SEP-1999;	9905-0152363.
PR 13-SEP-1999;	9905-0153070.
PR 15-SEP-1999;	9905-0153758.
PR 16-SEP-1999;	9905-0154018.
PR 20-SEP-1999;	9905-0154039.
PR 22-SEP-1999;	9905-0154779.
PR 23-SEP-1999;	9905-0155139.
PR 24-SEP-1999;	9905-0155486.
PR 28-SEP-1999;	9905-0155659.
PR 29-SEP-1999;	9905-0156458.
PR 04-OCT-1999;	9905-0156596.
PR 05-OCT-1999;	9905-0157117.
PR 06-OCT-1999;	9905-0157753.
PR 07-OCT-1999;	9905-0157865.
PR 08-OCT-1999;	9905-0158029.
PR 12-OCT-1999;	9905-0158232.
PR 13-OCT-1999;	9905-0158369.
PR 13-OCT-1999;	9905-0159293.
PR 13-OCT-1999;	9905-0159294.
PR 14-OCT-1999;	9905-0159329.
PR 14-OCT-1999;	9905-0159330.
PR 14-OCT-1999;	9905-0159331.
PR 14-OCT-1999;	9905-0159637.
PR 18-OCT-1999;	9905-0159684.
PR 21-OCT-1999;	9905-0159584.
PR 21-OCT-1999;	9905-0160741.
PR 21-OCT-1999;	9905-0160767.
PR 21-OCT-1999;	9905-0160768.
PR 21-OCT-1999;	9905-0160770.
PR 21-OCT-1999;	9905-0160814.
PR 21-OCT-1999;	9905-0160815.
PR 22-OCT-1999;	9905-0160980.
PR 22-OCT-1999;	9905-0160981.
PR 25-OCT-1999;	9905-0160989.
PR 25-OCT-1999;	9905-0161404.
PR 25-OCT-1999;	9905-0161405.
PR 26-OCT-1999;	9905-0161406.
PR 26-OCT-1999;	9905-0161359.
PR 26-OCT-1999;	9905-0161360.
PR 26-OCT-1999;	9905-0161361.
PR 28-OCT-1999;	9905-0161920.
PR 28-OCT-1999;	9905-0161922.
PR 28-OCT-1999;	9905-0161993.
PR 29-OCT-1999;	9905-0162142.
PR 29-OCT-1999;	9905-0162143.
PR 29-OCT-1999;	9905-0162228.
PR 01-NOV-1999;	9905-0162891.
PR 01-NOV-1999;	9905-0162894.
PR 01-NOV-1999;	9905-0162895.
PR 02-NOV-1999;	9905-0163081.
PR 02-NOV-1999;	9905-0163092.
PR 03-NOV-1999;	9905-0163248.
PR 03-NOV-1999;	9905-0163249.
PR 04-NOV-1999;	9905-0163281.
PR 04-NOV-1999;	9905-0163379.
PR 04-NOV-1999;	9905-0163380.
PR 08-NOV-1999;	9905-0163381.
PR 08-NOV-1999;	9905-0164146.

PR 08-NOV-1999; 99US-0164150.
 PR 08-NOV-1999; 99US-0164151.
 PR 09-NOV-1999; 99US-0164259.
 PR 09-NOV-1999; 99US-0164260.
 PR 10-NOV-1999; 99US-0164317.
 PR 10-NOV-1999; 99US-0164318.
 PR 10-NOV-1999; 99US-0164319.
 PR 10-NOV-1999; 99US-0164321.
 PR 10-NOV-1999; 99US-0164344.
 PR 10-NOV-1999; 99US-0164545.
 PR 10-NOV-1999; 99US-0164548.
 PR 12-NOV-1999; 99US-0164870.
 PR 12-NOV-1999; 99US-0164871.
 PR 12-NOV-1999; 99US-0164959.
 PR 12-NOV-1999; 99US-0164960.
 PR 12-NOV-1999; 99US-0164961.
 PR 12-NOV-1999; 99US-0164962.
 PR 15-NOV-1999; 99US-0164926.
 PR 15-NOV-1999; 99US-0164927.
 PR 15-NOV-1999; 99US-0164929.
 PR 16-NOV-1999; 99US-0165661.
 PR 16-NOV-1999; 99US-0165669.
 PR 16-NOV-1999; 99US-0165671.
 PR 17-NOV-1999; 99US-0165911.
 PR 17-NOV-1999; 99US-0165918.
 PR 17-NOV-1999; 99US-0165919.
 PR 18-NOV-1999; 99US-0166157.
 PR 18-NOV-1999; 99US-0166158.
 PR 18-NOV-1999; 99US-0166173.
 PR 19-NOV-1999; 99US-0166411.
 PR 19-NOV-1999; 99US-0166412.
 PR 19-NOV-1999; 99US-0166419.
 PR 22-NOV-1999; 99US-0166733.
 PR 22-NOV-1999; 99US-0166750.
 PR 23-NOV-1999; 99US-0167362.
 PR 24-NOV-1999; 99US-0167233.
 PR 24-NOV-1999; 99US-0167234.
 PR 24-NOV-1999; 99US-0167235.
 PR 24-NOV-1999; 99US-0167382.
 PR 30-NOV-1999; 99US-0167902.
 PR 30-NOV-1999; 99US-0167904.
 PR 30-NOV-1999; 99US-0167908.
 PR 01-DEC-1999; 99US-0168231.
 PR 01-DEC-1999; 99US-0168232.
 PR 01-DEC-1999; 99US-0168233.
 PR 02-DEC-1999; 99US-0168346.
 PR 02-DEC-1999; 99US-0168348.
 PR 02-DEC-1999; 99US-0168549.
 PR 03-DEC-1999; 99US-0168673.
 PR 03-DEC-1999; 99US-0168674.
 PR 03-DEC-1999; 99US-0168675.
 PR 07-DEC-1999; 99US-0169278.
 PR 07-DEC-1999; 99US-0169298.
 PR 07-DEC-1999; 99US-0169302.
 PR 08-DEC-1999; 99US-0169691.
 PR 08-DEC-1999; 99US-0169692.
 PR 16-DEC-1999; 99US-0171098.
 PR 16-DEC-1999; 99US-0171107.
 PR 16-DEC-1999; 99US-0171114.
 PR 19-JAN-2000; 2000US-0176867.
 PR 19-JAN-2000; 2000US-0176910.
 PR 26-JAN-2000; 2000US-0178166.
 PR 27-JAN-2000; 2000US-0177666.
 PR 27-JAN-2000; 2000US-0178344.
 PR 27-JAN-2000; 2000US-0178545.
 PR 27-JAN-2000; 2000US-0178546.
 PR 27-JAN-2000; 2000US-0178547.
 PR 28-JAN-2000; 2000US-0178754.
 PR 28-JAN-2000; 2000US-0178755.
 PR 01-FEB-2000; 2000US-0179385.
 PR 01-FEB-2000; 2000US-0179398.
 PR 03-FEB-2000; 2000US-0180039.

PR 03-FEB-2000; 2000US-0180139.
 PR 04-FEB-2000; 2000US-0180206.
 PR 04-FEB-2000; 2000US-0180207.
 PR 07-FEB-2000; 2000US-0180695.
 PR 07-FEB-2000; 2000US-0180696.
 PR 09-FEB-2000; 2000US-0181214.
 PR 09-FEB-2000; 2000US-0181228.
 PR 10-FEB-2000; 2000US-0181476.
 PR 10-FEB-2000; 2000US-0181551.
 PR 15-FEB-2000; 2000US-0182477.
 PR 15-FEB-2000; 2000US-0182478.
 PR 15-FEB-2000; 2000US-0182512.
 PR 15-FEB-2000; 2000US-0182516.
 PR 17-FEB-2000; 2000US-0183165.
 PR 17-FEB-2000; 2000US-0183166.
 XX
 PA (CERE-) CERES INC.
 PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
 PI Zheng L, Dumas J;
 XX
 DR WPI; 2000-507395/46.
 DR N-PSDB; AAC45095.
 XX
 PT New sequence determined DNA fragments (SDPs) from different plant
 PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
 PT protein coding sequences, untranslated regions, or as 3' termination
 PT sequences -
 XX
 PS Claim 19; SEQ ID 45283; 344pp + CD-ROM; English.
 CC The present sequence is a putative protein fragment from
 CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
 CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
 CC library which could then be sequenced, allowing the putative protein
 CC sequence(s) to be obtained. This sequence may be useful for protein
 CC identification and for aiding in the elucidation of signal transduction
 CC and metabolic pathways. Its coding sequence has a use in the control of
 CC gene expression as a promoter, coding sequence, 3'UTR or termination
 CC sequence, for controlling the behaviour of a gene within the chromosome,
 CC as a tool for use in genetic mapping, including a use in hybridisation
 CC assays, for recognition or isolation of similar DNA fragments, or for
 CC the identification of a particular organism.
 XX
 SQ Sequence 318 AA;
 AAG36898 Length: 318 May 30, 2002 09:25 Type: P Check: 464 ..
 1 MTPVGLGSPF TAAVSSSSP TTSSSTAVAVA DVTAMVSSSE EDLSKIRKP
 51 YTTKSRSEW TEPEHDKFLE ALQFDRDWK KIEAFISKT VIQIRSHAQ
 101 YFLKVRKSGT GEHLPPRPK RKAHPYPQK AHKNVQLOVP GSPKSTSEPN
 151 DPSFMFRPES SSMIMTSPTT AAAAWTNNNA OTISFTPLPK AYCCSTLMET
 201 AGAGANNNCS SSSSETPRRP SNRDARDHGN VGHSLRVLPD FAQVYFISGS
 251 VFDPYASNHL QKLRKMPID VETVILLMRN LSLINSSPDF EDHMTLDYK
 301 KRCIQGISNL LQKKNMKT
 IIAA_SEQUENCE 1.0
 ID AAG36899 standard; Protein; 284 AA.
 XX
 AC AAG36899;
 XX
 DF 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 45284.
 XX
 KW Protein identification: signal transduction pathway; metabolic pathway;
 hybridisation assay; genetic mapping; gene expression control; promoter;

KW	termination sequence.	
XX		
OS	Arabidopsis thaliana.	
XX	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-0301439.	
XX		
PR	25-FEB-1999; 99US-0121825.	PR 28-JUN-1999; 99US-0140823.
PR	05-MAR-1999; 99US-0121180.	PR 29-JUN-1999; 99US-0140991.
PR	09-MAR-1999; 99US-0123548.	PR 30-JUN-1999; 99US-0141287.
PR	23-MAR-1999; 99US-0125788.	PR 01-JUL-1999; 99US-0141842.
PR	25-MAR-1999; 99US-0126264.	PR 01-JUL-1999; 99US-0142154.
PR	29-MAR-1999; 99US-0126785.	PR 02-JUL-1999; 99US-0142055.
PR	01-APR-1999; 99US-0127462.	PR 06-JUL-1999; 99US-0142390.
PR	06-APR-1999; 99US-0128233.	PR 08-JUL-1999; 99US-0142803.
PR	08-APR-1999; 99US-0128714.	PR 09-JUL-1999; 99US-0142920.
PR	16-APR-1999; 99US-0129845.	PR 12-JUL-1999; 99US-0142977.
PR	19-APR-1999; 99US-0130077.	PR 13-JUL-1999; 99US-0143542.
PR	21-APR-1999; 99US-0130449.	PR 14-JUL-1999; 99US-0143624.
PR	23-APR-1999; 99US-0130510.	PR 15-JUL-1999; 99US-0144005.
PR	28-APR-1999; 99US-0130891.	PR 16-JUL-1999; 99US-0144085.
PR	30-APR-1999; 99US-0131449.	PR 16-JUL-1999; 99US-0144086.
PR	30-APR-1999; 99US-0132048.	PR 19-JUL-1999; 99US-0144325.
PR	04-MAY-1999; 99US-0132484.	PR 19-JUL-1999; 99US-0144335.
PR	05-MAY-1999; 99US-0132485.	PR 20-JUL-1999; 99US-0144352.
PR	06-MAY-1999; 99US-0132486.	PR 20-JUL-1999; 99US-0144632.
PR	06-MAY-1999; 99US-0132487.	PR 20-JUL-1999; 99US-0144884.
PR	07-MAY-1999; 99US-0132863.	PR 21-JUL-1999; 99US-0144884.
PR	11-MAY-1999; 99US-0134256.	PR 21-JUL-1999; 99US-0145086.
PR	14-MAY-1999; 99US-0134218.	PR 21-JUL-1999; 99US-0145088.
PR	14-MAY-1999; 99US-0134219.	PR 22-JUL-1999; 99US-0145085.
PR	14-MAY-1999; 99US-0134221.	PR 22-JUL-1999; 99US-0145087.
PR	14-MAY-1999; 99US-0134370.	PR 22-JUL-1999; 99US-0145089.
PR	18-MAY-1999; 99US-0134768.	PR 22-JUL-1999; 99US-0145192.
PR	19-MAY-1999; 99US-0134941.	PR 23-JUL-1999; 99US-0145145.
PR	20-MAY-1999; 99US-0135124.	PR 23-JUL-1999; 99US-0145218.
PR	21-MAY-1999; 99US-0135353.	PR 23-JUL-1999; 99US-0145224.
PR	24-MAY-1999; 99US-0135629.	PR 26-JUL-1999; 99US-0145276.
PR	25-MAY-1999; 99US-0136021.	PR 27-JUL-1999; 99US-0145913.
PR	27-MAY-1999; 99US-0136392.	PR 27-JUL-1999; 99US-0145918.
PR	28-MAY-1999; 99US-0136782.	PR 28-JUL-1999; 99US-0145951.
PR	01-JUN-1999; 99US-0137222.	PR 02-AUG-1999; 99US-0146386.
PR	03-JUN-1999; 99US-0137528.	PR 02-AUG-1999; 99US-0146388.
PR	04-JUN-1999; 99US-0137502.	PR 02-AUG-1999; 99US-0146389.
PR	07-JUN-1999; 99US-0137724.	PR 03-AUG-1999; 99US-0147038.
PR	08-JUN-1999; 99US-0138094.	PR 04-AUG-1999; 99US-0147204.
PR	10-JUN-1999; 99US-0138540.	PR 04-AUG-1999; 99US-0147302.
PR	10-JUN-1999; 99US-0138847.	PR 05-AUG-1999; 99US-0147192.
PR	14-JUN-1999; 99US-0138119.	PR 05-AUG-1999; 99US-0147260.
PR	16-JUN-1999; 99US-0139452.	PR 06-AUG-1999; 99US-0147303.
PR	16-JUN-1999; 99US-0139453.	PR 06-AUG-1999; 99US-0147416.
PR	17-JUN-1999; 99US-0139492.	PR 09-AUG-1999; 99US-0147493.
PR	18-JUN-1999; 99US-0139454.	PR 09-AUG-1999; 99US-0147935.
PR	18-JUN-1999; 99US-0139455.	PR 10-AUG-1999; 99US-0148177.
PR	18-JUN-1999; 99US-0139456.	PR 11-AUG-1999; 99US-0148319.
PR	18-JUN-1999; 99US-0139457.	PR 12-AUG-1999; 99US-0148341.
PR	18-JUN-1999; 99US-0139458.	PR 13-AUG-1999; 99US-0148565.
PR	18-JUN-1999; 99US-0139459.	PR 13-AUG-1999; 99US-0148684.
PR	18-JUN-1999; 99US-0139460.	PR 16-AUG-1999; 99US-0149368.
PR	18-JUN-1999; 99US-0139461.	PR 17-AUG-1999; 99US-0149175.
PR	18-JUN-1999; 99US-0139462.	PR 18-AUG-1999; 99US-0149426.
PR	18-JUN-1999; 99US-0139463.	PR 20-AUG-1999; 99US-0149722.
PR	18-JUN-1999; 99US-0139750.	PR 20-AUG-1999; 99US-0149723.
PR	18-JUN-1999; 99US-0139763.	PR 20-AUG-1999; 99US-0149920.
PR	21-JUN-1999; 99US-0139817.	PR 23-AUG-1999; 99US-0149902.
PR	22-JUN-1999; 99US-0139899.	PR 23-AUG-1999; 99US-0149930.
PR	23-JUN-1999; 99US-0140353.	PR 25-AUG-1999; 99US-0150566.
PR	23-JUN-1999; 99US-0140354.	PR 26-AUG-1999; 99US-0150884.
PR	24-JUN-1999; 99US-0140695.	PR 27-AUG-1999; 99US-0151065.
		PR 27-AUG-1999; 99US-0151066.
		PR 27-AUG-1999; 99US-0151080.
		PR 30-AUG-1999; 99US-0151303.
		PR 31-AUG-1999; 99US-0151438.
		PR 01-SEP-1999; 99US-0151930.
		PR 07-SEP-1999; 99US-0152363.

PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156556.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.
 PR 29-OCT-1999; 99US-0162143.
 PR 29-OCT-1999; 99US-0162228.
 PR 01-NOV-1999; 99US-0162891.
 PR 01-NOV-1999; 99US-0162894.
 PR 01-NOV-1999; 99US-0162895.
 PR 02-NOV-1999; 99US-0163091.
 PR 02-NOV-1999; 99US-0163092.
 PR 02-NOV-1999; 99US-0163093.
 PR 03-NOV-1999; 99US-0163248.
 PR 03-NOV-1999; 99US-0163249.
 PR 03-NOV-1999; 99US-0163281.
 PR 04-NOV-1999; 99US-0163379.
 PR 04-NOV-1999; 99US-0163380.
 PR 04-NOV-1999; 99US-0163381.
 PR 08-NOV-1999; 99US-0164146.
 PR 08-NOV-1999; 99US-0164150.
 PR 08-NOV-1999; 99US-0164151.
 PR 09-NOV-1999; 99US-0164259.
 PR 09-NOV-1999; 99US-0164260.
 PR 10-NOV-1999; 99US-0164317.
 PR 10-NOV-1999; 99US-0164318.
 PR 10-NOV-1999; 99US-0164319.
 PR 10-NOV-1999; 99US-0164321.
 PR 10-NOV-1999; 99US-0164544.
 PR 10-NOV-1999; 99US-0164545.
 PR 10-NOV-1999; 99US-0164548.
 PR 12-NOV-1999; 99US-0164870.
 PR 12-NOV-1999; 99US-0164871.
 PR 12-NOV-1999; 99US-0164959.

PR 12-NOV-1999; 99US-0164960.
 PR 12-NOV-1999; 99US-0164961.
 PR 12-NOV-1999; 99US-0164962.
 PR 15-NOV-1999; 99US-0164926.
 PR 15-NOV-1999; 99US-0164927.
 PR 15-NOV-1999; 99US-0164929.
 PR 16-NOV-1999; 99US-0165661.
 PR 16-NOV-1999; 99US-0165669.
 PR 16-NOV-1999; 99US-0165671.
 PR 17-NOV-1999; 99US-0165911.
 PR 17-NOV-1999; 99US-0165918.
 PR 17-NOV-1999; 99US-0165919.
 PR 18-NOV-1999; 99US-0166157.
 PR 18-NOV-1999; 99US-0166158.
 PR 18-NOV-1999; 99US-0166173.
 PR 19-NOV-1999; 99US-0166411.
 PR 19-NOV-1999; 99US-0166412.
 PR 19-NOV-1999; 99US-0166419.
 PR 22-NOV-1999; 99US-0166733.
 PR 22-NOV-1999; 99US-0166730.
 PR 23-NOV-1999; 99US-0167362.
 PR 24-NOV-1999; 99US-0167233.
 PR 24-NOV-1999; 99US-0167234.
 PR 24-NOV-1999; 99US-0167235.
 PR 24-NOV-1999; 99US-0167382.
 PR 30-NOV-1999; 99US-0167902.
 PR 30-NOV-1999; 99US-0167904.
 PR 01-DEC-1999; 99US-0168231.
 PR 01-DEC-1999; 99US-0168232.
 PR 01-DEC-1999; 99US-0168233.
 PR 02-DEC-1999; 99US-0168346.
 PR 02-DEC-1999; 99US-0168347.
 PR 02-DEC-1999; 99US-0168548.
 PR 03-DEC-1999; 99US-0168549.
 PR 03-DEC-1999; 99US-0168673.
 PR 03-DEC-1999; 99US-0168674.
 PR 07-DEC-1999; 99US-0168675.
 PR 07-DEC-1999; 99US-0169278.
 PR 07-DEC-1999; 99US-0169298.
 PR 07-DEC-1999; 99US-0169302.
 PR 08-DEC-1999; 99US-0169691.
 PR 08-DEC-1999; 99US-0169692.
 PR 16-DEC-1999; 99US-0171098.
 PR 16-DEC-1999; 99US-0171107.
 PR 16-DEC-1999; 99US-0171114.
 PR 19-JAN-2000; 2000US-0176866.
 PR 19-JAN-2000; 2000US-0176867.
 PR 19-JAN-2000; 2000US-0176910.
 PR 26-JAN-2000; 2000US-0178166.
 PR 27-JAN-2000; 2000US-0177666.
 PR 27-JAN-2000; 2000US-0178544.
 PR 27-JAN-2000; 2000US-0178545.
 PR 27-JAN-2000; 2000US-0178546.
 PR 27-JAN-2000; 2000US-0178547.
 PR 28-JAN-2000; 2000US-0178754.
 PR 28-JAN-2000; 2000US-0178755.
 PR 01-FEB-2000; 2000US-0179388.
 PR 01-FEB-2000; 2000US-0179395.
 PR 03-FEB-2000; 2000US-0180039.
 PR 03-FEB-2000; 2000US-0180139.
 PR 04-FEB-2000; 2000US-0180206.
 PR 04-FEB-2000; 2000US-0180207.
 PR 07-FEB-2000; 2000US-0180695.
 PR 07-FEB-2000; 2000US-0180696.
 PR 09-FEB-2000; 2000US-0181214.
 PR 09-FEB-2000; 2000US-0181228.
 PR 10-FEB-2000; 2000US-0181476.
 PR 10-FEB-2000; 2000US-0181551.
 PR 15-FEB-2000; 2000US-0182477.
 PR 15-FEB-2000; 2000US-0182478.
 PR 15-FEB-2000; 2000US-0182512.
 PR 15-FEB-2000; 2000US-0182516.
 PR 17-FEB-2000; 2000US-0183165.

PR 17-FEB-2000; 2000US-0183166.
XX
PA (CERE-) CERES INC.
XX
PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX
DR WPI; 2000-507395/46.
XX
N-PSDB; AAC45095.
PT
PT New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -
XX
PS Claim 19; SEQ ID 45284; 344pp + CD-ROM; English.
XX
XX The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3' UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.
XX
XX Sequence 284 AA;
SQ
AAG36899 Length: 284 May 30, 2002 09:25 Type: P Check: 6630 ..
1 MVSSSEEDLS KIRKPYTIT KRSRWTEPE HDKPIEALQL FDRRWKITEA
51 FIGSKTVIQL RSHAKYFLK VOKSGTEHL PPPRRKRAA HPPYOKAHKN
101 VOQVPGSFK STSEPNDSF MFRPESSML WTSPTTAAA PWTNNAQTIS
151 FTPLPKAVCC STLMETAGAG ANNCCSSSE NTPRRSRND ARDGNVGHG
201 LRVLPDFAOV YGFISVFDP YASNHLQKLK KMDPIDVETV LLAMRNLSIN
251 LSPDFEDHW HTLDYKKRCI OGISNLDQK NMKT
11AA_SEQUENCE 1.0
ID AAG37272 standard; Protein; 584 AA.
XX
AC AAG37272;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 45804.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR

PR 19-NOV-1999; 99US-0166412.
 PR 19-NOV-1999; 99US-0166419.
 PR 22-NOV-1999; 99US-0166713.
 PR 22-NOV-1999; 99US-0166750.
 PR 23-NOV-1999; 99US-0167362.
 PR 24-NOV-1999; 99US-0167334.
 PR 24-NOV-1999; 99US-0167233.
 PR 24-NOV-1999; 99US-0167235.
 PR 24-NOV-1999; 99US-0167382.
 PR 30-NOV-1999; 99US-0167902.
 PR 30-NOV-1999; 99US-0167904.
 PR 30-NOV-1999; 99US-0167908.
 PR 01-DEC-1999; 99US-0168231.
 PR 01-DEC-1999; 99US-0168232.
 PR 01-DEC-1999; 99US-0168233.
 PR 02-DEC-1999; 99US-0168546.
 PR 02-DEC-1999; 99US-0168548.
 PR 02-DEC-1999; 99US-0168549.
 PR 03-DEC-1999; 99US-0168673.
 PR 03-DEC-1999; 99US-0168674.
 PR 03-DEC-1999; 99US-0168675.
 PR 07-DEC-1999; 99US-0169278.
 PR 07-DEC-1999; 99US-0169298.
 PR 08-DEC-1999; 99US-0169302.
 PR 08-DEC-1999; 99US-0169691.
 PR 16-DEC-1999; 99US-0169692.
 PR 16-DEC-1999; 99US-0171098.
 PR 16-DEC-1999; 99US-0171107.
 PR 16-DEC-1999; 99US-0171114.
 PR 19-JAN-2000; 2000US-0176666.
 PR 19-JAN-2000; 2000US-0176667.
 PR 19-JAN-2000; 2000US-0176867.
 PR 26-JAN-2000; 2000US-0178160.
 PR 27-JAN-2000; 2000US-0177666.
 PR 27-JAN-2000; 2000US-0178544.
 PR 27-JAN-2000; 2000US-0178545.
 PR 27-JAN-2000; 2000US-0178546.
 PR 27-JAN-2000; 2000US-0178547.
 PR 28-JAN-2000; 2000US-0178754.
 PR 28-JAN-2000; 2000US-0178755.
 PR 01-FEB-2000; 2000US-0179388.
 PR 01-FEB-2000; 2000US-0179395.
 PR 03-FEB-2000; 2000US-0180039.
 PR 03-FEB-2000; 2000US-0180139.
 PR 04-FEB-2000; 2000US-0180206.
 PR 04-FEB-2000; 2000US-0180207.
 PR 07-FEB-2000; 2000US-0180695.
 PR 07-FEB-2000; 2000US-0180696.
 PR 09-FEB-2000; 2000US-0181214.
 PR 10-FEB-2000; 2000US-0181228.
 PR 10-FEB-2000; 2000US-0181476.
 PR 10-FEB-2000; 2000US-0181551.
 PR 15-FEB-2000; 2000US-0182477.
 PR 15-FEB-2000; 2000US-0182478.
 PR 15-FEB-2000; 2000US-0182512.
 PR 15-FEB-2000; 2000US-0182516.
 PR 17-FEB-2000; 2000US-0183165.
 PR 17-FEB-2000; 2000US-0183166.
 XX (CERE-) CERES INC.
 PA Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
 PI zheng L, Dumas J;
 XX WPI; 2000-507395/46.
 DR N-PSDB; AAC45242.
 XX
 XX New sequence determined DNA fragments (SDFs) from different plant
 PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters
 PT protein coding sequences, untranslated regions, or as 3' termination
 PT sequences -
 XX
 XX Claim 19; SEQ ID 45804; 344pp + CD-ROM; English.

XX		The present sequence is a putative protein fragment from
CC		Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC		Rt-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC		library which could then be sequenced, allowing the putative protein
CC		sequence(s) to be obtained. This sequence may be useful for protein
CC		identification and for aiding in the elucidation of signal transduction
CC		and metabolic pathways. Its coding sequence has a use in the control of
CC		gene expression as a promoter, coding sequence, 3'UTR or termination
CC		sequence, for controlling the behaviour of a gene within the chromosome,
CC		as a tool for use in genetic mapping, including a use in hybridisation
CC		assays, for recognition or isolation of similar DNA fragments, or for
CC		the identification of a particular organism.
SQ	Sequence	584 AA;
AAG37272	Length:	584 May 30, 2002 09:25 Type: P Check: 7852 ..
1	YFEIIVLLFL LIRIRONSSF KLEFINLVEEF WSEIYMGSGS VTSIADGFRF	
51	HPTDELVRY YLKRIKCNKP FKFDALSYVD YKSEPMDLP DKSRLKSRDL	
101	EYFFFSMLDK KYRNGSKTNK AREMGWTKT GKDEKLNGS KVGGMKKTLV	
151	YHKGRAPRGE RTNKMWEHYR LYDQDIKDTG VHQAFAVLGR IFQKSGSGPK	
201	NQEQRGAPFV EEEWEEDDM TEVPDDLDLG SEDHYYVMD DIDQSENFY	
251	VYDAIPPIPLN FIGHESSNV ETNYSDSTNY IQQGNTYKDS GGFEPQPAES	
301	YEKDQKPIIR DRDGLONEG ICGGVADKHS ETLQSDNIF GTDTCYNDF	
351	PVESNYVLGE AFLDPNSNL ENDGLYLEFN DLSTQQDGF DFEDYLTFPD	
401	EFFDPSQLMG NEDVFEPDOE LRQEVYTKEL EKKEISRKH VEEKKEDEA	
451	SCSKQVDADA TEEPDPKYRP LTRKASHMUG AIPAPLANAS EFFTQDAIR	
501	LHAAGSSGSV HVPAGMITTS DSMNGMSYCK NENIDLILSL GLVOGNTPAE	
551	KSGNSSANAM LIEMCFVILL LSYSFKVSIL VSSR	
IIAA_SEQUENCE 1.0		
ID	AAG37273 standard; Protein: 549 AA.	
XX	AAG37273;	
AC		
DT	18-OCT-2000 (first entry)	
XX		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 45805.	
XX		
KM	Protein identification: signal transduction pathway; metabolic pathway;	
KM	hybridisation assay; genetic mapping; gene expression control; promoter	
term	termination sequence.	
OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
PF	25-FEB-2000; 2000EP-0301439.	
XX		
PR	25-FEB-1999; 99US-0121825.	
PR	05-MAR-1999; 99US-0123160.	
PR	09-MAR-1999; 99US-0123548.	
PR	23-MAR-1999; 99US-0125788.	
PR	25-MAR-1999; 99US-0126264.	
PR	29-MAR-1999; 99US-0126785.	
PR	01-APR-1999; 99US-0127462.	
PR	06-APR-1999; 99US-0128234.	
PR	08-APR-1999; 99US-0128714.	
PR	16-APR-1999; 99US-0129845.	

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-APR-1999; 99US-0132407.
PR 05-APR-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164116.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.

PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 03-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169279.
PR 07-DEC-1999; 99US-0169288.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.
XX (CERE-) CERES INC.
PA Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX WPI; 2000-507395/46.
DR N-PSDB; AAC45242.
XX
XX New sequence determined DNA fragments (SDFs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
XX sequences -
XX
PS Claim 19; SEQ ID 45805; 344pp + CD-ROM; English.
XX
CC The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA

CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.

XX Sequence 549 AA;

AG37273 Length: 549 May 30, 2002 09:25 Type: P Check: 1025 ..

1 MGKSVTSLA PGRFRHPIDE ELVRYTLRK ICNPKFFDA ISVTDYKSE
51 PWDLPKSRSL KSRDLEWYFF SMLDKYRNG SKTRNATEMG YMKTGKDRE
101 ILNGSKVYGM KKTIVYHKGR APRGERTMWV MHEKRLVDD LDKTGVHDA
151 FYLCRIFFOKS GSGPKNGEY GAPVEEEME EEDDMTFVPD QEDLGSEDHV
201 YVHMDIDQK SENFVYDAI PIPLFINGE SSNNVETNS DSINYIQGTG
251 NWMDSGGYE QPAEYKQD KPIIRDQGS LQNGIGCGV QDKHSTLOS
301 SDNIFGTDS CYNDPPEVSN YLIGFAFLDP NSNLENDGL YLETNDLSST
351 QDGFDEFDY LTFEDEFDP SOLMGNEHYF FDOELFQEV ETEKEKEET
401 SRSKHVVEK EKDEASCSKO VDADATEREP DYKPYLAKA SHHGLAIRAP
451 LANASEPPT DAIIRLHAAQ SSGSVHTAG MITSDSNMG WSYGKNENID
501 LILSLGLVYG NTAPEKSGNS SAMMILFMC FWLLLSVSF KVSILVSR
11AA_SEQUENCE 1.0
ID AG37274 standard; Protein: 478 AA.
XX
AC AAG37274;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 45806.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 99US-0132048.

PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135323.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 05-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159658.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164344.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165671.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166733.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.

PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.

PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155466.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 18-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 22-OCT-1999; 9905-0160989.

```

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161923.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 02-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 03-NOV-1999; 99US-0163348.
PR 03-NOV-1999; 99US-0163349.
PR 03-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 12-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0167362.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.

```

```

PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176867.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181224.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 15-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.

```

(CERE-) CERES INC.

PA Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;

DR MPI; 2000-507395/46.

XX N-PSDB; AAC45684.

PT New sequence determined DNA fragments (SDPs) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -

PS Claim 19; SEQ ID 47387; 344pp + CD-ROM; English.

XX The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.

XX Sequence 261 AA;

AAG38413 Length: 261 May 30, 2002 09:25 Type: P Check: 3958 ..

1 MWOSOSISTL TICGSVKVSS LIRNRLNSYK ASSLIGDRCV SCQFLKSPS
51 FRSHWKSIAQ RNLRLVEARW PFQGGEGQGL DPSSERSESA NEDILIFFQ
101 LDLATRQVYA MNLEQYDIAQ QUREKITEVE EESTIRLOGK RGSASKEAQ
151 DKGISIRLR ADIQNAIDSE DYGLAKLRD EISKLEASL AVSAKALAFE
201 KAEYAFRLQK KLRHKTFGYR AVVCGMDPIC SESSSMWMA EVELPRGSN
251 QPFYQLKTI C
11AA_SEQUENCE 1.0
ID MAG38839 standard; Protein; 986 AA.
XX
AC AAG38839;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47973.
XX
KM Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999;

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146399.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147319.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.

PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 10-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164870.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166730.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 02-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169278.
PR 07-DEC-1999; 99US-0169278.
PR 08-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 16-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.

PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179389.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.

XX (CERE-) CERES INC.

PA Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;

PI Zheng L, Dumas J;

DR WPI; 2000-507395/46.

XX N-PSDB; AAC45844.

XX New sequence determined DNA fragments (SDFS) from different plant
PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
PT protein coding sequences, untranslated regions, or as 3' termination
PT sequences -

XX Claim 19; SEQ ID 47973; 344pp + CD-ROM; English.

XX The present sequence is a putative protein fragment from
CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
CC library which could then be sequenced, allowing the putative protein
CC sequence(s) to be obtained. This sequence may be useful for protein
CC identification and for aiding in the elucidation of signal transduction
CC and metabolic pathways. Its coding sequence has a use in the control of
CC gene expression as a promoter, coding sequence, 3'UTR or termination
CC sequence, for controlling the behaviour of a gene within the chromosome,
CC as a tool for use in genetic mapping, including a use in hybridisation
CC assays, for recognition or isolation of similar DNA fragments, or for
CC the identification of a particular organism.

XX Sequence 986 AA;

AAG38839 Length: 986 May 30, 2002 09:25 Type: P Check: 7690 ..

1 MGRSVAVVT APPSATAPGT ASVATSLAPG FRHPTDEEL VSYLKRKVL
51 GQPYRPAIG EVDIYKHEPW DIAVSRUKT RDQEWYFSA LDKRYNGAR
101 MNRATNGYW KATGKDREIR RDILLGMKK TLVHSGRAP DGLRTVMVMH
151 EYRLVEYTE KNGNLVQDAY VLGRVFIKNN IGPPSGRNYA PRMEEEHADD
201 EGALIPIDV KLRLEPPVA NGNDOMDEI OSASKSLINI NEPPRETAPL
251 DIESDOONH ENDLKPEHN NNNYDENEE TLKREOMEER ERPPRCVYL
301 NKEAPLPLQ YKRRROSEN NNSRNTQDH CSSTTTVDN TTLLISSAA
351 ATNTAISALL EFSLMGISDK KEKPOOPLRP HKEPLPQTP LASPEERKS

401 VCYEIIIVL FLIRIRONS SEKLFINIV EFWSIEYMG RGSYSLAPGF
451 RHPTDEELV RYLLKRKCN KFKKPAISV TDVYKSEPD LPDKSRKSR
501 DLEWTFMSL DKRYRSGKT NRATEGYWK TTGKDREILN GSKVYNGKKT
551 LVYHKGRAIR GERTNWMHE YRLVDODLCK TGVHODAEVL CRIFOKSGSG
601 PKNEQYGAP FVEEWEED DMTFVPDQED LGSEDHVYVH MDDIDOKSEN
651 FVYTDALPIP LNFIGHSSN NVETNSDSI NYIQOTGYW DSGGYFEQPA
701 ESYEKDQPI IRDRDGLON EGICGVQDK HSEFLQSSDN IFGNDTSCYN
751 DFPVSNALI GEAFIDPNSN LLENDGLYIE TNDLSSTQD GFPEEYLTF
801 FDEFDPQSL MGNEDVFFDQ EELFOVEYETK ELEKEETSRS KHVEEREKD
851 EASCSKVDA DATEFEPDYK YPLKKASHM LGAIAPPLAN ASEPTKDA
901 IRLAAQSSG SVHVTAGMIT IDSNNMGWSY GKNEMLDIL SLGLVGNATA
951 PEKSGNSSAW AMLIFMCFW LLLSVSKVYS ILVSSR

IIA SEQUENCE 1.0
ID AAG38840 standard; Protein; 886 AA.

XX AAG38840;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 47974.

XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130891.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.

PR 19-MAY-1999;	99US-0134941.	PR 28-JUL-1999;	99US-0145951.
PR 20-MAY-1999;	99US-0135124.	PR 02-AUG-1999;	99US-0146386.
PR 21-MAY-1999;	99US-0135353.	PR 02-AUG-1999;	99US-0146388.
PR 24-MAY-1999;	99US-0135629.	PR 03-AUG-1999;	99US-0146389.
PR 25-MAY-1999;	99US-0136021.	PR 03-AUG-1999;	99US-0147038.
PR 27-MAY-1999;	99US-0136392.	PR 04-AUG-1999;	99US-0147204.
PR 28-MAY-1999;	99US-0136782.	PR 04-AUG-1999;	99US-0147302.
PR 01-JUN-1999;	99US-0137222.	PR 05-AUG-1999;	99US-0147192.
PR 03-JUN-1999;	99US-0137528.	PR 05-AUG-1999;	99US-0147260.
PR 04-JUN-1999;	99US-0137502.	PR 06-AUG-1999;	99US-0147303.
PR 07-JUN-1999;	99US-0137724.	PR 06-AUG-1999;	99US-0147416.
PR 08-JUN-1999;	99US-0138094.	PR 09-AUG-1999;	99US-0147493.
PR 10-JUN-1999;	99US-0138540.	PR 09-AUG-1999;	99US-0147935.
PR 10-JUN-1999;	99US-0138847.	PR 10-AUG-1999;	99US-0148171.
PR 14-JUN-1999;	99US-0139119.	PR 11-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139452.	PR 12-AUG-1999;	99US-0148341.
PR 16-JUN-1999;	99US-0139453.	PR 13-AUG-1999;	99US-0148565.
PR 17-JUN-1999;	99US-0139492.	PR 13-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139454.	PR 16-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0139899.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140353.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140354.	PR 01-SEP-1999;	99US-0151930.
PR 24-JUN-1999;	99US-0140695.	PR 07-SEP-1999;	99US-0152363.
PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 29-JUN-1999;	99US-0140991.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154018.
PR 01-JUL-1999;	99US-0141842.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154479.
PR 02-JUL-1999;	99US-0142055.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142820.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144684.	PR 18-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0144506.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145376.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.

```

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161362.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.
PR 29-OCT-1999; 99US-0162228.
PR 01-NOV-1999; 99US-0162891.
PR 01-NOV-1999; 99US-0162894.
PR 01-NOV-1999; 99US-0162895.
PR 02-NOV-1999; 99US-0163091.
PR 02-NOV-1999; 99US-0163092.
PR 02-NOV-1999; 99US-0163093.
PR 03-NOV-1999; 99US-0163248.
PR 03-NOV-1999; 99US-0163249.
PR 03-NOV-1999; 99US-0163281.
PR 04-NOV-1999; 99US-0163379.
PR 04-NOV-1999; 99US-0163380.
PR 04-NOV-1999; 99US-0163381.
PR 08-NOV-1999; 99US-0164146.
PR 08-NOV-1999; 99US-0164150.
PR 08-NOV-1999; 99US-0164151.
PR 09-NOV-1999; 99US-0164259.
PR 09-NOV-1999; 99US-0164260.
PR 10-NOV-1999; 99US-0164317.
PR 10-NOV-1999; 99US-0164318.
PR 10-NOV-1999; 99US-0164319.
PR 10-NOV-1999; 99US-0164321.
PR 10-NOV-1999; 99US-0164544.
PR 10-NOV-1999; 99US-0164545.
PR 12-NOV-1999; 99US-0164548.
PR 12-NOV-1999; 99US-0164871.
PR 12-NOV-1999; 99US-0164959.
PR 12-NOV-1999; 99US-0164960.
PR 12-NOV-1999; 99US-0164961.
PR 12-NOV-1999; 99US-0164962.
PR 15-NOV-1999; 99US-0164926.
PR 15-NOV-1999; 99US-0164927.
PR 15-NOV-1999; 99US-0164929.
PR 16-NOV-1999; 99US-0165661.
PR 16-NOV-1999; 99US-0165669.
PR 16-NOV-1999; 99US-0165671.
PR 17-NOV-1999; 99US-0165911.
PR 17-NOV-1999; 99US-0165918.
PR 17-NOV-1999; 99US-0165919.
PR 18-NOV-1999; 99US-0166157.
PR 18-NOV-1999; 99US-0166158.
PR 18-NOV-1999; 99US-0166173.
PR 19-NOV-1999; 99US-0166411.
PR 19-NOV-1999; 99US-0166412.
PR 19-NOV-1999; 99US-0166419.
PR 22-NOV-1999; 99US-0166733.
PR 22-NOV-1999; 99US-0166750.
PR 23-NOV-1999; 99US-0167362.
PR 24-NOV-1999; 99US-0167233.
PR 24-NOV-1999; 99US-0167234.
PR 24-NOV-1999; 99US-0167235.
PR 24-NOV-1999; 99US-0167382.
PR 30-NOV-1999; 99US-0167902.
PR 30-NOV-1999; 99US-0167904.
PR 30-NOV-1999; 99US-0167908.
PR 01-DEC-1999; 99US-0168231.
PR 01-DEC-1999; 99US-0168232.
PR 01-DEC-1999; 99US-0168233.
PR 02-DEC-1999; 99US-0168546.
PR 02-DEC-1999; 99US-0168548.
PR 03-DEC-1999; 99US-0168549.
PR 03-DEC-1999; 99US-0168673.
PR 03-DEC-1999; 99US-0168674.
PR 03-DEC-1999; 99US-0168675.
PR 07-DEC-1999; 99US-0169278.

```

```

PR 07-DEC-1999; 99US-0169298.
PR 07-DEC-1999; 99US-0169302.
PR 08-DEC-1999; 99US-0169691.
PR 08-DEC-1999; 99US-0169692.
PR 16-DEC-1999; 99US-0171098.
PR 16-DEC-1999; 99US-0171107.
PR 16-DEC-1999; 99US-0171114.
PR 19-JAN-2000; 2000US-0176866.
PR 19-JAN-2000; 2000US-0176867.
PR 19-JAN-2000; 2000US-0176910.
PR 26-JAN-2000; 2000US-0178166.
PR 27-JAN-2000; 2000US-0177666.
PR 27-JAN-2000; 2000US-0178544.
PR 27-JAN-2000; 2000US-0178545.
PR 27-JAN-2000; 2000US-0178546.
PR 27-JAN-2000; 2000US-0178547.
PR 28-JAN-2000; 2000US-0178754.
PR 28-JAN-2000; 2000US-0178755.
PR 01-FEB-2000; 2000US-0179388.
PR 01-FEB-2000; 2000US-0179395.
PR 03-FEB-2000; 2000US-0180039.
PR 03-FEB-2000; 2000US-0180139.
PR 04-FEB-2000; 2000US-0180206.
PR 04-FEB-2000; 2000US-0180207.
PR 07-FEB-2000; 2000US-0180695.
PR 07-FEB-2000; 2000US-0180696.
PR 09-FEB-2000; 2000US-0181214.
PR 09-FEB-2000; 2000US-0181228.
PR 10-FEB-2000; 2000US-0181476.
PR 10-FEB-2000; 2000US-0181551.
PR 15-FEB-2000; 2000US-0182477.
PR 15-FEB-2000; 2000US-0182478.
PR 15-FEB-2000; 2000US-0182512.
PR 15-FEB-2000; 2000US-0182516.
PR 17-FEB-2000; 2000US-0183165.
PR 17-FEB-2000; 2000US-0183166.

XX (CERE-) CERES INC.
XX Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L, Dumas J;
XX WPI: 2000-507395/46.
XX N-PSDB; AAC45844.
XX
XX New sequence determined DNA fragments (SDFs) from different plant
XX species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
XX protein coding sequences, untranslated regions, or as 3' termination
XX sequences -
XX
XX Claim 19; SEQ ID 47974; 344pp + CD-ROM; English.
XX
XX The present sequence is a putative protein fragment from
XX Arabidopsis thaliana. Its coding sequence was isolated by carrying out
XX RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
XX library which could then be sequenced, allowing the putative protein
XX sequence(s) to be obtained. This sequence may be useful for protein
XX identification and for aiding in the elucidation of signal transduction
XX and metabolic pathways. Its coding sequence has a use in the control of
XX gene expression as a promoter, coding sequence, 3' UTR or termination
XX sequence, for controlling the behaviour of a gene within the chromosome,
XX as a tool for use in genetic mapping, including a use in hybridisation
XX assays, for recognition or isolation of similar DNA fragments, or for
XX the identification of a particular organism.
XX
XX Sequence 886 AA:
XX
XX

```

AAG38840 Length: 886 May 30, 2002 09:25 Type: P Check: 7810 ..

```

1 MNRATNRGTYW KATGKREIR RDILLGMMK TLVPHSGRAP DGLRTNWMH
51 EYRLVEYTE KGNLVODAY VLGRVPHKN ICPPSGNRYA PFMEBEMADD

```

101 EGALIPGIDV KLRLEPPVA NGNDQMDDEI QSASKSLINI NEPPREAPL
151 DIESDQNNH ENDLKPEHN NNNNYDENBE TLKRQMEEE ERPPRCVL
201 NKEAPLPLLO YKRRROSEN NNSRRNTQDH CSSTTTVDN TTTLISSAA
251 ATNNAISML EFSLMGISDK KKKPOOPLRP HKEPLPPOTP LASEEERKS
301 VCPEEITVL FLIRIRQNS SEKLEINLIV EFWSEIYNGR GSVTSLAPG
351 RFHPTDEELV RYLLKRRICN KPFEKDAISV TDVYKSEPMO LPDKSRILKS
401 DLEWYFFSML DKRYRNGSKT NRATEMGYWK TTGKDRLELN GSKYVGNKKT
451 LVYHKGAPR GERTNYMWE YALVQDIDDK TGVHODAFVL CRIFQKSGSG
501 PKNGEQYCAP FVEEWEEDD DMTFVPDQED LGSEDHVYVH MDDIQSEEN
551 FVVYDAIPIPI LNFIGHESN NVEITNSDSI NYIQOTGNM DSGGYEQPA
601 ESYEKDOKPI IRDRDGLQNL EGICGGVODK HSETLQSSDN IFGTDTSCYN
651 DFPVESNVLJ GEAFILPNSN LLENDGLYLE TNDLSSFOOD GFDPEDYLTG
701 FDEFEDDSQL MGNEDVFEFD EELFOVEVEIK ELEKETRSR KHYVEEKED
751 EASCSKQVDA DATEFEDPYK YPLLKASHM LGAIPIAPLAN ASEPTKDDAA
801 IRLHAOSSG SVHVTAGMIT ISDSNMGWGY GKNEHLILIL SLGIQGNFTA
851 PERGNSNSAW AMLIFMCFWV LLTSLSEKVS ILVSR

!1AA_SEQUENCE 1.0
ID AAG38841 standard; Protein; 859 AA.
XX AAG38841;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47975.
XX
KM Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay: genetic mapping; gene expression control; promoter;
KM termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132047.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139451.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145951.
PR 28-JUL-1999;	99US-0145986.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155133.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157353.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
PR 21-OCT-1999;	99US-0160815.
PR 22-OCT-1999;	99US-0160980.
PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160989.
PR 25-OCT-1999;	99US-0161404.
PR 25-OCT-1999;	99US-0161405.
PR 25-OCT-1999;	99US-0161406.
PR 26-OCT-1999;	99US-0161359.
PR 26-OCT-1999;	99US-0161360.
PR 26-OCT-1999;	99US-0161361.
PR 28-OCT-1999;	99US-0161920.
PR 28-OCT-1999;	99US-0161992.
PR 28-OCT-1999;	99US-0161993.
PR 29-OCT-1999;	99US-0162142.
PR 29-OCT-1999;	99US-0162143.
PR 29-OCT-1999;	99US-0162228.
PR 01-NOV-1999;	99US-0162891.
PR 01-NOV-1999;	99US-0162894.
PR 01-NOV-1999;	99US-0162895.
PR 02-NOV-1999;	99US-0163091.
PR 02-NOV-1999;	99US-0163092.
PR 03-NOV-1999;	99US-0163093.
PR 03-NOV-1999;	99US-0163248.
PR 03-NOV-1999;	99US-0163249.
PR 03-NOV-1999;	99US-0163281.
PR 04-NOV-1999;	99US-0163379.
PR 04-NOV-1999;	99US-0163380.
PR 08-NOV-1999;	99US-0164146.
PR 08-NOV-1999;	99US-0164150.
PR 08-NOV-1999;	99US-0164151.
PR 09-NOV-1999;	99US-0164259.
PR 09-NOV-1999;	99US-0164260.
PR 10-NOV-1999;	99US-0164317.
PR 10-NOV-1999;	99US-0164318.
PR 10-NOV-1999;	99US-0164319.
PR 10-NOV-1999;	99US-0164321.
PR 10-NOV-1999;	99US-0164544.
PR 10-NOV-1999;	99US-0164545.
PR 12-NOV-1999;	99US-0164962.
PR 12-NOV-1999;	99US-0164962.
PR 15-NOV-1999;	99US-0164926.
PR 15-NOV-1999;	99US-0164927.
PR 15-NOV-1999;	99US-0164929.
PR 16-NOV-1999;	99US-0165661.
PR 16-NOV-1999;	99US-0165669.
PR 16-NOV-1999;	99US-0165671.
PR 17-NOV-1999;	99US-0165911.
PR 17-NOV-1999;	99US-0165918.
PR 17-NOV-1999;	99US-0165919.
PR 18-NOV-1999;	99US-0166157.
PR 18-NOV-1999;	99US-0166158.
PR 18-NOV-1999;	99US-0166173.
PR 19-NOV-1999;	99US-0166411.
PR 19-NOV-1999;	99US-0166412.
PR 19-NOV-1999;	99US-0166419.
PR 22-NOV-1999;	99US-0166733.
PR 22-NOV-1999;	99US-0166750.
PR 23-NOV-1999;	99US-0167362.
PR 24-NOV-1999;	99US-0167233.
PR 24-NOV-1999;	99US-0167234.
PR 24-NOV-1999;	99US-0167235.
PR 30-NOV-1999;	99US-0167382.
PR 30-NOV-1999;	99US-0167902.
PR 30-NOV-1999;	99US-0167904.
PR 30-NOV-1999;	99US-0167908.
PR 01-DEC-1999;	99US-0168231.
PR 01-DEC-1999;	99US-0168232.

PR 01-DEC-1999: 99US-0168233.
 PR 02-DEC-1999: 99US-0168546.
 PR 02-DEC-1999: 99US-0168548.
 PR 02-DEC-1999: 99US-0168549.
 PR 03-DEC-1999: 99US-0168573.
 PR 03-DEC-1999: 99US-0168574.
 PR 03-DEC-1999: 99US-0168675.
 PR 07-DEC-1999: 99US-0169278.
 PR 07-DEC-1999: 99US-0169298.
 PR 07-DEC-1999: 99US-0169302.
 PR 08-DEC-1999: 99US-0169691.
 PR 08-DEC-1999: 99US-0169692.
 PR 16-DEC-1999: 99US-0171098.
 PR 16-DEC-1999: 99US-0171107.
 PR 16-DEC-1999: 99US-0171114.
 PR 19-JAN-2000: 2000US-0176866.
 PR 19-JAN-2000: 2000US-0176867.
 PR 19-JAN-2000: 2000US-0176910.
 PR 26-JAN-2000: 2000US-0178166.
 PR 27-JAN-2000: 2000US-0177666.
 PR 27-JAN-2000: 2000US-0178544.
 PR 27-JAN-2000: 2000US-0178545.
 PR 27-JAN-2000: 2000US-0178546.
 PR 27-JAN-2000: 2000US-0178547.
 PR 28-JAN-2000: 2000US-0178754.
 PR 28-JAN-2000: 2000US-0178755.
 PR 01-FEB-2000: 2000US-0179388.
 PR 01-FEB-2000: 2000US-0179395.
 PR 03-FEB-2000: 2000US-0180039.
 PR 03-FEB-2000: 2000US-0180139.
 PR 04-FEB-2000: 2000US-0180206.
 PR 04-FEB-2000: 2000US-0180207.
 PR 07-FEB-2000: 2000US-0180695.
 PR 07-FEB-2000: 2000US-0180696.
 PR 09-FEB-2000: 2000US-0181214.
 PR 09-FEB-2000: 2000US-0181228.
 PR 10-FEB-2000: 2000US-0181476.
 PR 10-FEB-2000: 2000US-0181551.
 PR 15-FEB-2000: 2000US-0182477.
 PR 15-FEB-2000: 2000US-0182478.
 PR 15-FEB-2000: 2000US-0182512.
 PR 15-FEB-2000: 2000US-0182516.
 PR 17-FEB-2000: 2000US-0183165.
 PR 17-FEB-2000: 2000US-0183166.
 XX (CERE-) CERES INC.
 PA Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
 PI Zheng L, Dumas J;
 XX MPI: 2000-507395/46.
 DR N-PSDB; AAC45844.
 XX
 PT New sequence determined DNA fragments (SDFs) from different plant
 PT species, e.g. corn, rice or Arabidopsis thaliana, useful as promoters,
 PT protein coding sequences, untranslated regions, or as 3' termination
 PT sequences -
 XX
 PS Claim 19; SEQ ID 47975; 344pp + CD-ROM; English.
 XX
 CC The present sequence is a putative protein fragment from
 CC Arabidopsis thaliana. Its coding sequence was isolated by carrying out
 CC RT-PCR on all of the mRNA obtained from the plant, and creating a cDNA
 CC library which could then be sequenced, allowing the putative protein
 CC sequence(s) to be obtained. This sequence may be useful for protein
 CC identification and for aiding in the elucidation of signal transduction
 CC and metabolic pathways. Its coding sequence has a use in the control of
 CC gene expression as a promoter, coding sequence, 3'UTR or termination
 CC sequence, for controlling the behaviour of a gene within the chromosome,
 CC as a tool for use in genetic mapping, including a use in hybridisation
 CC assays, for recognition or isolation of similar DNA fragments, or for
 CC the identification of a particular organism.
 XX

SQ Sequence 859 AA;
 AAG38841 Length: 859 May 30, 2002 09:25 Type: P Check: 4753 ..
 1 MKKTLVFRHS RAPDGLRTNW VMHEKRLVEY ETEKNGNLVQ DAYLVCVFI
 51 KNIIGPPSGN RYAPFMEEM ADDEGALING IDVKLRLEPP PVANGDMD
 101 QELISASKSL ININEPPRET APDIESDQO NHHENDLKPE EHNNNNNYDE
 151 NEETLKREOM EEEERPPRPV CYLNKEAPLP LLQYKRRQS ESNNNSSRNT
 201 QDHCSSTTTT VDNITTLISS SAAATNTAIS ALLEFSLMGI SDKKEKQOP
 251 LRPHKEPLPP QTPPLASPEEK RKVCYFEIIL VLLFLLIRIR QNSSFKLEIN
 301 LVFEFWSLEY MGRGSVTSIA PGRFRHPIDE ELVRYLLARK ICKNPFEDA
 351 ISVTDVYKSE PMDLPDKSL KSRDLEWTFE SMLDKKIRNG SKTNRAEMG
 401 YWKTGKDRRE ILNGSKVGM KKTLYYHKGR APRGERTMW MHEKRLVDQD
 451 LDKTGVHQA FVLCRIFFQKS GSGPKNGEQ GAPFVEEME EDDMTFVPD
 501 QEDIGSEDPV YVHMDIDQK SENEVYVDAL PIPLNFHGE SSNNVETNYS
 551 DSINYIOOTG NYMDSGYFE QPAESYEKQD KPIIRDRGS LONEGICGV
 601 QDKHSETLOS SDNIFGHTS CYNDPVESEN YLIGRAFIDP NSNLEMDGL
 651 YLETNDLSST QODGFDEYD LTFEDETDP SOLMGNEVDF FDEELFQEV
 701 ETKELKEKET SRSKHVVEEK EKDEASCQKQ VDADATEFEP DYKYPILKKA
 751 SHMLGAIPAP LANASEPTK DAIRLHAQQ SSGSVHYTAG MITSOSNMG
 801 WSYKNEMLD LIISLGIYQ NTAPKSGNS SAMANLIMC FWVLLSYST
 851 KVSILVSSR
 !!AA-SEQUENCE 1.0
 ID AAY95439 standard; Protein: 1503 AA.
 XX
 AC AAY95439;
 XX
 DT 10-OCT-2000 (first entry)
 XX
 DE Human calcium channel polypeptide.
 XX
 KW Human: SOC-2/CRAC-1; calcium channel; store operated channel;
 KW calcium release activated channel; therapy; diagnosis;
 KW lymphocyte proliferative disorder.
 XX
 OS Homo sapiens.
 OS
 PN W0200040614-A2.
 XX
 PD 13-JUL-2000.
 XX
 PF 20-DEC-1999: 99WO-US29996.
 XX
 PR 30-DEC-1998: 98US-0114220.
 PR 29-JAN-1999: 99US-0120018.
 PR 22-JUN-1999: 99US-0140415.
 XX
 PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
 XX
 PI Scharenberg AM;
 XX
 DR MPI: 2000-465957/40.
 DR N-PSDB; AAA49926.
 XX
 PT New SOC/CRAC calcium channel polynucleotides and polypeptides used to

PT diagnose and treat proliferative disorders associated with the channel,
PT and to screen for novel modulators of the channel .
PS
XX Example; Page 69-72; 108pp; English.
XX
CC The present sequence is that of a human calcium channel polypeptide
CC as deduced from a cDNA clone (see AAA4926) identified in an EST
CC database search for sequences showing homology to Caenorhabditis
CC elegans calcium signal sequences. Clones isolated from such
CC screenings were used to identify human clones (see AAA4922-24)
CC encoding members (see AA95435-37) of a new family of SOC (store
CC operated channel) or CRAC (calcium release activated channel)
CC calcium channel polypeptides. SOC/CRAC polypeptides modulate Ca2+
CC flux into and out of a cell, and may be activated upon depletion
CC of Ca2+ from intracellular calcium stores, allowing Ca2+ influx
CC into a cell. SOC/CRAC polypeptides and polynucleotides can
CC be used to treat patients that require modulation of calcium influx
CC into their SOC/CRAC-expressing cells. Polypeptides can be used to
CC deliver therapeutic and/or imaging agents to such cells to modulate
CC proliferation and growth. SOC/CRAC polypeptides also represent
CC targets for designing and/or identifying inhibitors that block
CC lymphocyte proliferation and binding agents that selectively bind
CC to SOC/CRAC polypeptides to which drugs or toxins can be conjugated
CC for delivery to SOC/CRAC expressing cells.
XX
SQ Sequence 1503 AA;

AA95439 Length: 1503 May 30, 2002 09:25 Type: P Check: 5055 ..

1 MEPALRKAG SEQEGFEG LPRVTDLGW SNLRSSNSL FSKWRLCCPF
51 GNNDKOSLS SWIPENIKK ECVYEVESK LSDAGKVCQ CGYHEOHL
101 EATKPHFG TQMDPKHVO EMPDAFGDI VFTGISOVK KYAVSODTP
151 SSVYHLMATQ HMGLDVFNLL ISVTGAKNF NKKPRLKSI F RGLVKAQT
201 TGAWITGGS HTGVKMGVE AVRDESLSSS YKEGELITIG VATWGVHRR
251 EGLHPHGSF PAEYILDEG QGNLCLDSN HSHFLVDDG THQGYVEIP
301 LRTLEKFI S EOTKRGVA IKIPIVCYL EGGPGLTHTI DNATNGTSPC
351 VVEGSGRVA DVAQVANLP VSDITISLIQ OKLSYFQEM FETFTESRIY
401 EMTKKIDIV RRQQLTVFR ECKDQGDVD VAILQALLKA SRSDHGHGE
451 NMDHQLLAY AMNRVDIARS ETFMDEMOK PSDLHPMTA ALISNKEFV
501 KLFLNGVOL KEFVTWDTLL YLYENLDPSG LFSKIQKVL VEDERPACA
551 PAARLQWNN VAQVLRRLG DFTQPLPRP RHNDRLRLLL PVPHVKLNVO
601 GVSRLSLYKR SSGHVTFTMD PIRDLITMAI VONRRELGI IMAOSQDCA
651 AALACSKILK ELKSEEDTD SSEEMALAE EYEHRAIGVF TECYRKDER
701 AQKLLTRVSE AMGKTICLOL ALEAKDMKFV SHGIGQALFT KYWKGQLSVD
751 NGLRWITICM LARPLLITGL ISERREKLOD VGTPARARA FFIAPVVVFH
801 LNLISYFAFL CLFAYVLMDV FQVPFSWCEC AIYLMFSLV CEEMROLFYD
851 PDEGLMKMKA ALVESDFWNK LDVGATILFV AGLTCRLIPA TLVGRVILS
901 LDFILFCLRL MHFTTISKTL GKRIIYVKRM MKDVEFFLEL LAVWVSFV
951 AKQAILIINE RYVDWLFEGA VYHSYLTIFG QIPGYIDGN FNPEHCSPNG
1001 TDPKPKCPE SDATQORPAF PEMLVYLLIC LYLFNTNILL LNLIAFENV
1051 TFOOVQEHNT QIMKFORHDL IEHYHGRPAA PPPFILLSHL QLFIRYVVLK

1101 TPARKHOKL NLEKNEEA LLSWEIYLKE NYLONROFOQ KORPEOKIED
1151 ISNKVDAMVD LLDLDPLKRS GSWEQRLASL EQVNAQTFARA LHWIVFTLRA
1201 SGFSSEADVP TLASQKAEE PDAEPGCRKK TEEPDSITHV NARHLIYPMC
1251 PVTREPVNE KVPETEFLI YDPPFYTABR KDAAAMPDG DTLEPLSTIQ
1301 YNVVDGLNDR RSFNGPYTVO AGLPLNPMGR TGLRGSGLS CFQPHHTLYP
1351 MYTRRRNED GAIQCKSIKK MLEVLYVKLP LSEHMLPCG SREPGEMLPK
1401 KLKRLROEH WPSFENDLKC GMEVYKGYMD DFRNTDNAMI ETVAVSVHFQ
1451 QDNDELNRL NSNNHACDSG ASIRMQVDR RPLVANKHT LLOKAAAEFG
1501 AHY
11AA-SEQUENCE 1.0
ID AA66728 standard; protein: 280 AA.
XX
AC AA66728;
XX
DT 05-APR-2000 (first entry)
XX
DE Membrane-bound protein PRO1131.
XX
KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
KW pharmaceutical; receptor Immunoadhesin; gene mapping.
XX
OS Homo sapiens.
XX
PN W0963088-A2.
XX
PD 09-DEC-1999.
XX
PE 02-JUN-1999; 99WO-US12252.
XX
PR 02-JUN-1998; 98US-0087607.
PR 02-JUN-1998; 98US-0087609.
PR 02-JUN-1998; 98US-0087759.
PR 03-JUN-1998; 98US-0087827.
PR 04-JUN-1998; 98US-0088021.
PR 04-JUN-1998; 98US-0088025.
PR 04-JUN-1998; 98US-0088028.
PR 04-JUN-1998; 98US-0088029.
PR 04-JUN-1998; 98US-0088030.
PR 04-JUN-1998; 98US-0088033.
PR 04-JUN-1998; 98US-0088326.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088212.
PR 05-JUN-1998; 98US-0088217.
PR 05-JUN-1998; 98US-0088655.
PR 10-JUN-1998; 98US-0088730.
PR 10-JUN-1998; 98US-0088734.
PR 10-JUN-1998; 98US-0088738.
PR 10-JUN-1998; 98US-0088740.
PR 10-JUN-1998; 98US-0088741.
PR 10-JUN-1998; 98US-0088742.
PR 10-JUN-1998; 98US-0088810.
PR 10-JUN-1998; 98US-0088811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 10-JUN-1998; 98US-0088826.
PR 11-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089090.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.

PR 16-JUN-1998; 98US-0089512.
PR 16-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089598.
PR 17-JUN-1998; 98US-0089599.
PR 17-JUN-1998; 98US-0089600.
PR 17-JUN-1998; 98US-0089603.
PR 18-JUN-1998; 98US-0089607.
PR 18-JUN-1998; 98US-0089608.
PR 19-JUN-1998; 98US-0089647.
PR 19-JUN-1998; 98US-0089648.
PR 19-JUN-1998; 98US-0089652.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 24-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 25-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090576.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090682.
PR 26-JUN-1998; 98US-0090683.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091344.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 07-JUL-1998; 98US-0091982.
PR 09-JUL-1998; 98US-0092182.
PR 10-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 11-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.

PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
PR 17-AUG-1998; 98US-0096891.
PR 17-AUG-1998; 98US-0096894.
PR 17-AUG-1998; 98US-0096895.
PR 17-AUG-1998; 98US-0096897.
PR 18-AUG-1998; 98US-0096949.
PR 18-AUG-1998; 98US-0096950.
PR 18-AUG-1998; 98US-0096959.
PR 18-AUG-1998; 98US-0096960.
PR 18-AUG-1998; 98US-0097022.
PR 19-AUG-1998; 98US-0097141.
PR 20-AUG-1998; 98US-0097218.
PR 24-AUG-1998; 98US-0097611.
PR 26-AUG-1998; 98US-0097951.
PR 26-AUG-1998; 98US-0097952.
PR 26-AUG-1998; 98US-0097954.
PR 26-AUG-1998; 98US-0097955.
PR 26-AUG-1998; 98US-0097957.
PR 26-AUG-1998; 98US-0097971.
PR 26-AUG-1998; 98US-0097974.
PR 26-AUG-1998; 98US-0097978.
PR 26-AUG-1998; 98US-0097979.
PR 26-AUG-1998; 98US-0097986.
PR 26-AUG-1998; 98US-0098014.
PR 31-AUG-1998; 98US-0098525.
PR 16-SEP-1998; 98US-0100634.
PR 12-JAN-1999; 99US-0115565.
XX
XX (GETH) GENENTECH INC.
XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
PI Wood WI, Yuan J;
PI
XX WPI: 2000-072883/06.
XX N-PSDB; AA265072.
XX
XX Membrane-bound proteins and related nucleotide sequences -
XX
XX claim 12; Fig 230; 822P; English.

XX The invention provides membrane-bound PRO polypeptides and
XX polynucleotides encoding them. The PRO sequences of the invention were
XX identified based on extracellular domain homology screening. The PRO
XX sequences have homology with proteins including LDL receptors, TIE
XX ligands and various enzymes. The membrane-bound proteins and receptor
XX molecules are useful as pharmaceutical and diagnostic agents. Receptor
XX immunoadhesins, for instance, can be used as therapeutic agents to block
XX receptor-ligand interactions. The membrane-bound proteins can also be
XX employed for screening of potential peptide or small molecule inhibitors
XX of the relevant receptor/ligand interaction. The PRO encoding sequences
XX are useful as hybridization probes, in chromosome and gene mapping and in
XX the generation of antisense RNA and DNA. PRO nucleic acid sequences
XX will also be useful for the preparation of PRO polypeptides, especially
XX by recombinant techniques.

XX Sequence 280 AA;

AA66728 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..

1 MQAKYSSTRD MDDDGDTTM SIHSQASATT RHPEPRTEH RABSSWRPV
51 ALTLTLCLIV LLIGLAALGL LFFQYQULSN TGDDPTISOME ERLGNTSOEL
101 OSLOVONIKL ACSLSHVAEK LCRELNNKAG AHRSPCTEQ WKKHGNCVQ
151 FYKDSKSWED CKYFLSENS TMLKINKQED LEFAASQSYS EFFYSYWTGL
201 LRPDGKRAWL MWDGPFPTSE LFHIIIDYTS PRSHCVAIL NGMIFSKDKC
251 ELKRCVCERR AGWVKPESLIH VPPEITLGHGD

IIAA_SEQUENCE 1.0

ID AAY76218 standard; Protein: 161 AA.
 XX
 AC AAY76218;
 XX
 DT 23-MAR-2000 (first entry)
 XX
 DE Human secreted protein encoded by gene 95.
 XX
 KW Human; secreted protein; cancer; tumour; developmental abnormality;
 KW fetal deficiency; blood disorder; immune system disorder; inflammation;
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
 KW therapy; chromosome 1.
 XX
 OS Homo sapiens.
 XX
 PN WO958660-A1.
 XX
 PD 18-NOV-1999.
 XX
 PF 06-MAY-1999; 99WO-US09847.
 XX
 PR 12-MAY-1998; 98US-0085093.
 PR 12-MAY-1998; 98US-0085094.
 PR 12-MAY-1998; 98US-0085105.
 PR 12-MAY-1998; 98US-0085180.
 PR 18-MAY-1998; 98US-0085906.
 PR 18-MAY-1998; 98US-0085920.
 PR 18-MAY-1998; 98US-0085921.
 PR 18-MAY-1998; 98US-0085922.
 PR 18-MAY-1998; 98US-0085923.
 PR 18-MAY-1998; 98US-0085928.
 PR 18-MAY-1998; 98US-0085928.
 PR 18-MAY-1998; 98US-0085925.
 PR 18-MAY-1998; 98US-0085927.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
 PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR;
 PI Lafleur DM, Endress GA, Ebner R;
 XX
 DR WPI: 2000-062296/05.
 DR N-PSDB: AA265344.
 XX
 PS Claim 11; Page 412; 475pp; English.
 XX
 CC AA265350 to AA265350 represent 97 isolated human secreted protein genes.
 CC AAY76124 to AAY76223 are the secreted proteins encoded by the 97 human
 CC genes. The gene encoding this protein was found to be on chromosome 1.
 CC The genes and their corresponding secreted polypeptides are
 CC useful for preventing, treating or ameliorating medical conditions,
 CC e.g. by protein or gene therapy. Also pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new genes. Specific
 CC uses are described for each of the 97 genes, based on which tissues they
 CC are most highly expressed in, and include developing products for the
 CC diagnosis or treatment of cancer, tumours, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
 CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
 CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
 CC disorders, digestive/endocrine disorders, infections and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.
 CC The sequences shown in AAY76224 to AAY76424 represent fragments of the
 CC secreted proteins.
 CC
 XX

SQ Sequence 161 AA;
 AAY76218 Length: 161 May 30, 2002 09:25 Type: P Check: 8440 ..
 1 MAFTPAFCY MSLIVLCAAL IFFPAIWHITA FDELFTDRKS PIDOCNPVHA
 51 RERLRNIERI CFLRLKLVLP EYSIHSLFCI MELCAQEWILT LGILNVPILFFY
 101 HFWRFHCPA DSESLAYDPP VVMNADTLSY CQKEXXCKLA FYLSFFYYIL
 151 YCMITTLVSS X
 IIAA_SEQUENCE 1.0
 ID AAY53783 standard; Protein: 541 AA.
 XX
 AC AAY53783;
 XX
 DT 22-FEB-2000 (first entry)
 XX
 DE A maize phosphate transporter designated Zmpt1.
 XX
 KW Phosphate transporter; Zmpt1; phosphorous uptake; phosphate poor soil;
 KW free phosphorus; animal consumption; seed; phosphate level; germination.
 XX
 OS Zea mays.
 XX
 PN WO958657-A2.
 XX
 PD 18-NOV-1999.
 XX
 PF 11-MAY-1999; 99WO-US10374.
 XX
 PR 13-MAY-1998; 98US-0085292.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI McElver JA, Zhao S, Bowen BA, Wang X;
 XX
 DR WPI: 2000-053096/04.
 DR N-PSDB: AA236416.
 XX
 PT Zea mays phosphate transporter genes and proteins, transgenic plants
 PT and seeds -
 XX
 PS Claim 9; Page 86-88; 108pp; English.
 XX
 CC The present sequence represents a Zea mays phosphate transporter,
 CC designated Zmpt1. The polynucleotide sequence was identified and isolated
 CC from a Zea mays nucleic acid library. The phosphate transporter
 CC polynucleotides and polypeptides are useful for altering phosphate
 CC transporter expression in transgenic plants. They are useful for
 CC providing plants capable of increasing phosphorous uptake in phosphate
 CC poor soils, or to increase the level of free phosphorus in plants used
 CC for animal consumption, or to decrease the level of free phosphate in
 CC seeds where phosphate levels would otherwise be deleterious to
 CC germination. The invention also provides means for identifying compounds
 CC that bind to and/or increase or decrease the enzymatic activity of the
 CC Zmpt polypeptides.
 XX
 SQ Sequence 541 AA;
 AAY53783 Length: 541 May 30, 2002 09:25 Type: P Check: 6982 ..
 1 MARGDGIQV LSLADAKTQ WHFTLIIVA GMGFTTDAVD LFCISLVTKL
 51 LGRYYTDTG KDNESLPPN VAAAVNGVAF CGTLAQLQLEF GWLGDKLGRK
 101 SVYGTMLLM VICSVASGLS FGHHTPGVMA TLCEFFRWLG FGIGDYPIS
 151 ATINSEYAK KTGCAFTAAV FAMQGRGILA GGIYLVIVSA AFRAGYAPA
 201 YRDHFNSTV PQADYVWRII LILGAAPAML TYWBMKME TARYTALVAK

251 NAKQAAADMS RVLTQETIVDE QEKLDQAVTA ESNTFGLEFSR EFARRRGLHL
 301 VGTSTTWELL DIAFYSONLF QKDIPTTSINW IPKANTMSAL EEEVRRISRAQ
 351 TLIALCGTVP GYWFVVALID VYGRFAIQLL GFPMVTVFML GLAIPIYHHW
 401 TTPGNHIGFV VMAFTFFEFAN FGPNSTTFTV PAEIFPARLR STCHGISAAS
 451 GKAGAIIGAF FGLYAAQND KSKADAGYPA GIGVRSLSLFV LAASNILGFI
 501 VLFVLPESKG KSLKEMSGEA DDAEDDAVGT RAVRPSGTQM V

11AA_SEQUENCE 1.0
 ID AAY53784 standard; Protein: 542 AA.

AC AAY53784;
 XX
 DT 22-FEB-2000 (first entry)

DE A maize phosphate transporter designated ZmptII.

XX Phosphate transporter; ZmptII; phosphorous uptake; phosphate poor soil;
 KW free phosphorus; animal consumption; seed; phosphate level; germination.
 XX

OS Zea mays.

PN WO958657-A2.

XX 18-NOV-1999.

PD 11-MAY-1999; 99WO-US10374.

XX 13-MAY-1998; 98US-0085292.

XX (PION-) PIONEER HI-BRED INT INC.

XX MCEliver JA, Zhao S, Bowen BA, Wang X;

PI WPI; 2000-053096/04.

DR N-PSDB; AA236417.

XX Zea mays phosphate transporter genes and proteins, transgenic plants

PT and seeds -

PS Claim 9; Page 90-91; 108pp; English.

XX The present sequence represents a Zea mays phosphate transporter,
 CC designated ZmptII. The polynucleotide sequence was identified and
 CC isolated from a Zea mays nucleic acid library. The phosphate transporter
 CC polynucleotides and polypeptides are useful for altering phosphate
 CC transporter expression in transgenic plants. They are useful for
 CC providing plants capable of increasing phosphorous uptake in phosphate
 CC poor soils, or to increase the level of free phosphorus in plants used
 CC for animal consumption, or to decrease the level of free phosphate in
 CC seeds where phosphate levels would otherwise be deleterious to
 CC germination. The invention also provides means for identifying compounds
 CC that bind to and/or increase or decrease the enzymatic activity of the
 CC Zmpt polypeptides.
 XX

XX Sequence 542 AA;

SO AAY53784 Length: 542 May 30, 2002 09:25 Type: P Check: 8403 ..

1 MARGDGLQV LSLADAKTO WHHTAIIYA CMGFETDAYD LFCISLVTKL
 51 LGRITTYDTS KDSPLSPN VAAAVNGYAF CGTLAQOLF GWLGDGLGRK
 101 SVYGMTLMMV VICSVASGLS FGHPTGYMA TLCCFFRWLG FGIIGDYPPLS
 151 ATIMSEYANK RRGAFIAAV FAMOGFGILA GGIYTLVISA AFRAAVPSPA
 201 YRDHFTSTV PQADFVWRYI LMLGAAPALL TTYWRMKMPE TARYTALVAK

251 NAKQAAADMS KVILTEILDE QEKLDQAVTA EGANSFGIFS REPARRHGLH
 301 LVGTATTWFL LDIAFYSONLF POKDIPTTSIN WIPKANTMSA LEEVRRISRA
 351 QTLIALCGTV PGYWFVVALI DVGGRFAIQL LGFPMVTVFV LGLAIPIYHHW
 401 TTPGNHIGFV VMAFTFFFEA NEGPNSTFTI VPAEIFPARL RSTCHGISAA
 451 SKAGAIIGA FGLYAAQND DSKTDAGYP AGIGVRSLSFV LAASNILGFI
 501 VLFVLPESKG KSLKEMSGEA AEDSEEPYG ARAVRPSGTQ MV

11AA_SEQUENCE 1.0
 ID AAY53785 standard; Protein: 547 AA.

AC AAY53785;
 XX
 DT 22-FEB-2000 (first entry)

DE A maize phosphate transporter designated ZmptIII.

XX Phosphate transporter; ZmptIII; phosphorous uptake; phosphate poor soil;
 KW free phosphorus; animal consumption; seed; phosphate level; germination.
 XX

OS Zea mays.

PN WO958657-A2.

XX 18-NOV-1999.

PD 11-MAY-1999; 99WO-US10374.

XX 13-MAY-1998; 98US-0085292.

XX (PION-) PIONEER HI-BRED INT INC.

XX MCEliver JA, Zhao S, Bowen BA, Wang X;

PI WPI; 2000-053096/04.

DR N-PSDB; AA238418.

XX Zea mays phosphate transporter genes and proteins, transgenic plants

PT and seeds -

PS Claim 9; Page 94-95; 108pp; English.

XX The present sequence represents a Zea mays phosphate transporter,
 CC designated ZmptIII. The polynucleotide sequence was identified and
 CC isolated from a Zea mays nucleic acid library. The phosphate transporter
 CC polynucleotides and polypeptides are useful for altering phosphate
 CC transporter expression in transgenic plants. They are useful for
 CC providing plants capable of increasing phosphorous uptake in phosphate
 CC poor soils, or to increase the level of free phosphorus in plants used
 CC for animal consumption, or to decrease the level of free phosphate in
 CC seeds where phosphate levels would otherwise be deleterious to
 CC germination. The invention also provides means for identifying compounds
 CC that bind to and/or increase or decrease the enzymatic activity of the
 CC Zmpt polypeptides.
 XX

XX Sequence 547 AA;

SO AAY53785 Length: 547 May 30, 2002 09:25 Type: P Check: 6498 ..

1 MAAGDLEVLJ ALDTAKTQW HETAIVAGM GFETDAYDLF CISLVTKLIG
 51 RIYTVGSA TPGTLRPHVS ASVNGVAFVG TLSQLTFGW LGLKGRKKV
 101 YGMTLMMVL CSVASGLSFG HTPASVMTL CFFRFWLGFG IGGDYPPLSAT
 151 IMSEYANKT RCAFIAVPA MGGFGIMAGC LVAIVSAMF KASFAPAYVA
 201 VDPAASTFPQ ADFVWRIIM LGAMPALTY YMKMKETA RYALVALAKNA

251 KQAAADSKV LOVEISAGAA EGGAAATAT ATEPAPASAS ASRGLFSGEF
 301 LRRHGLHLG TSLTWFLDI AFYSONLFQK DIFSAVGWIP KAAATMALKEE
 351 LFSIARRQSL IALGTVPGV WFTVALIDLV GRFAIQVTGF LMMIVFNLGL
 401 AVPEHHWTP GHHIGIVMT GLEFFANFG PNATTEIVPA EIPPARLST
 451 CHGISAASGK IGAIGSFQF LYLAONRQPA KTDHGYPAGI GVRNSLFLLA
 501 GCNLLGLAFT FLVPESKGKS LEMSGENDE AAAAATPNY NNRIVPV

11AA_SEQUENCE 1.0
 ID AAY53786 standard; Protein: 539 AA.

AC AAY53786;

DT 22-FEB-2000 (first entry)

DE A maize phosphate transporter designated Zmptv.

KW Phosphate transporter; zmptv; phosphorous uptake; phosphate poor soil;
 free phosphorus; animal consumption; seed; phosphate level; germination.

OS Zea mays.

PN MO958657-A2.

PD 18-NOV-1999.

PF 11-MAY-1999; 99WO-US10374.

PR 13-MAY-1998; 98US-0085292.

PA (PION-) PIONEER HI-BRED INT INC.

PI McElver JA, Zhao S, Bowen BA, Wang X;

DR WPI: 2000-053096/04.
 N-PSDB; AA236419.

PT Zea mays phosphate transporter genes and proteins, transgenic plants
 and seeds -

PS Claim 9; Page 98-99; 108bp; English.

CC The present sequence represents a Zea mays phosphate transporter,
 designated zmptv. The polynucleotide sequence was identified and
 isolated from a Zea mays nucleic acid library. The phosphate transporter
 polynucleotides and polypeptides are useful for altering phosphate
 transporter expression in transgenic plants. They are useful for
 providing plants capable of increasing phosphorous uptake in phosphate
 poor soils, or to increase the level of free phosphorus in plants used
 for animal consumption, or to decrease the level of free phosphate in
 seeds where phosphate levels would otherwise be deleterious to
 germination. The invention also provides means for identifying compounds
 that bind to and/or increase or decrease the enzymatic activity of the
 zmpt polypeptides.

CC Sequence 539 AA;

AA53786 Length: 539 May 30, 2002 09:25 Type: P Check: 3508 ..

1 MARGDGLQV LSLADAAKTQ WYHFAIIVA GMGFTDAYD LFCISLVTKL

51 LGRITYYDTS KDSPGSLPPN VAAAYNGVAF CGTLAQLQLEF GWLGDKGRK

101 SYVGMLTVM VICSVASGLS FGHPTGVMA TLCEFFEWLG FGIGGDYPLS

151 ATIMSEYANK RTGAFIAAV FAMQGFILA GGIYTLVISA AFRAAYPSPA

201 YRDHFISTV PQADIVRWYI VMLGAAPALL TTYWLMKPE TARTYALVAK

251 NAKQAAAADMS KVLHTEIVDE QEKLDAAEGA NSFGLFSREF ARRHGLHVG
 301 TATTWFLDI AFYSONLFQK DIFTSINWIP KANTMSALEE VYRISRAQTL
 351 IALGTVPGV WFTVALIDVY GRFAIOLLEF FMYVFMGL ALPHHWTTP
 401 GNHIGFVVMY ATEFFANFG PNSTTEIVPA EIPPARLST CHGISAASGK
 451 AGAIIAGARF LYAAQONDRS KTDAGYPAGI GVRNSLFLVLA ASNMLGFVLT
 501 FLVPESKGKS LEMSGEAD SEEPYGGARA VRPSETQWY

11AA_SEQUENCE 1.0
 ID AAY53787 standard; Protein: 509 AA.

AC AAY53787;

DT 22-FEB-2000 (first entry)

DE A maize phosphate transporter designated Zmptv.

KW Phosphate transporter; zmptv; phosphorous uptake; phosphate poor soil;
 free phosphorus; animal consumption; seed; phosphate level; germination.

OS Zea mays.

PN MO958657-A2.

PD 18-NOV-1999.

PF 11-MAY-1999; 99WO-US10374.

PR 13-MAY-1998; 98US-0085292.

PA (PION-) PIONEER HI-BRED INT INC.

PI McElver JA, Zhao S, Bowen BA, Wang X;

DR WPI: 2000-053096/04.
 N-PSDB; AA236420.

PT Zea mays phosphate transporter genes and proteins, transgenic plants
 and seeds -

PS Claim 9; Page 101-102; 108bp; English.

CC The present sequence represents a Zea mays phosphate transporter,
 designated zmptv. The polynucleotide sequence was identified and
 isolated from a Zea mays nucleic acid library. The phosphate transporter
 polynucleotides and polypeptides are useful for altering phosphate
 transporter expression in transgenic plants. They are useful for
 providing plants capable of increasing phosphorous uptake in phosphate
 poor soils, or to increase the level of free phosphorus in plants used
 for animal consumption, or to decrease the level of free phosphate in
 seeds where phosphate levels would otherwise be deleterious to
 germination. The invention also provides means for identifying compounds
 that bind to and/or increase or decrease the enzymatic activity of the
 zmpt polypeptides.

CC Sequence 509 AA;

AA53787 Length: 509 May 30, 2002 09:25 Type: P Check: 2070 ..

1 MALQOQLRVL HALDVARTQL YHFMATAIAG MGFTDAYDF FAISLVMDLI

51 SYLYNEQID RGVKATINGI ALCGAVPGOL VFGWLGRMG RKRIYGVTL

101 LMMVTSLASG LYEFTNEASN VVAVLCEFRF WLGFIGGDY PLSATIMSEY

151 ANKRRAQAFI AAVFAMQGF NLAAGIVAVY VSASLRTNP RRANFPWRI

201 VMLGANPAI LITYWRKMP ETARTALVA KDARRAASDM SSVLHVEIIP

251 EDNAVQDKY GLFSAQFLRY HGHILATSA CWLAVDTTFY SLNLYKKDIF
 301 ADVGLIDPPG NNDLFTRMVY TTLHTGTIAL CGTLPGYFFV VAFVDRIGRY
 351 RIQLGFTMM SVLTALIAAT YAWKROETI QRMKFAVLY GLTNFRANG
 401 PNTTTFVPA EIFPARMRAT CHGTAAGFK IGALIGVGF MSMEHGVV
 451 PRLMALFAS NLVGLVTFEL LPDSKGSLE EMAGTEEOQ QOQDAAVVA
 501 ADHINLVPI
 11AA_SEQUENCE 1.0
 ID AAY53788 standard; Protein; 554 AA.
 AC AAY53788;
 XX
 DT 22-FEB-2000 (first entry)
 DE A maize phosphate transporter designated ZmptVI.
 XX
 XX Phosphate transporter; ZmptVI; phosphorous uptake; phosphate poor soil;
 KW free phosphorus; animal consumption; seed; phosphate level; germination.
 XX
 OS Zea mays.
 XX
 PN WO958657-A2.
 PD 18-NOV-1999.
 XX
 PF 11-MAY-1999; 99MO-US10374.
 PR 13-MAY-1998; 98US-0085292.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 PI McElver JA, Zhao S, Bowen BA, Wang X;
 XX
 DR WPI; 2000-053096/04.
 DR N-PSDB; AA36421.
 XX
 PT Zea mays phosphate transporter genes and proteins, transgenic plants
 PT and seeds -
 XX
 PS Claim 9; Page 105-106; 108pp; English.
 XX
 CC The present sequence represents a Zea mays phosphate transporter,
 CC designated ZmptVI. The polynucleotide sequence was identified and
 CC isolated from a Zea mays nucleic acid library. The phosphate transporter
 CC polynucleotides and polypeptides are useful for altering phosphate
 CC transporter expression in transgenic plants. They are useful for
 CC providing plants capable of increasing phosphorous uptake in phosphate
 CC poor soils, or to increase the level of free phosphorus in plants used
 CC for animal consumption, or to decrease the level of free phosphate in
 CC seeds where phosphate levels would otherwise be deleterious to
 CC germination. The invention also provides means for identifying compounds
 CC that bind to and/or increase or decrease the enzymatic activity of the
 CC Zmpt polypeptides.
 CC
 XX Sequence 554 AA:
 SQ
 AAY53788 Length: 554 May 30, 2002 09:25 Type: P Check: 9398 ..

251 RYALLEGNA KQANDMQRY MDVEIQAEQD KLARYKAND YPLSREPAR
 301 RQGLHLIGTA TTMFLIDIAF YSQNLTKQDI FPAIKITSPV DINALKEVF
 351 EISKAMFLVA LGTFPGYVW TVALIDKNGR YLIQLIOFFM MSYEMLMQV
 401 MYNDLKNKHT TLPALFYALT FFFANFGPNS TTFVLPALFL PYRVSTCHA
 451 ISASCKRAGA IYAARFQVSL TLKGDVGHK KALILSVTN ILGFPTFLV
 501 PETMGRLSEE ISGEDGVNEN GPAPAGVAM GVADVSKDK MPVSTREMS
 551 SMHA
 11AA_SEQUENCE 1.0
 ID ABB63883 standard; Protein; 2040 AA.
 AC ABB63883;
 XX
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 18441.
 XX
 XX Drosophila melanogaster polypeptide SEQ ID NO 18441.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW
 OS Drosophila melanogaster.
 OS
 PN WO200171042-A2.
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001MO-US09231.
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABLO7986.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 18441; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABLO1840-ABLO16175) and the encoded proteins
 CC (ABBS7737-ABBS7072).
 CC
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 2040 AA:
 SQ
 ABB63883 Length: 2040 May 30, 2002 09:25 Type: P Check: 9003 ..

1 MAAPGGSNLA VIDALDSART QMYHMKALVI AGMGFTTDAY DLFCISTVSK
 51 LIGRIYPPD NLVIDKPRKG TLPVSNNMV TGVALVGTLM GOLVEGYFED
 101 KLGKRVYGI TLVMAACAI GSGLSFGSSA HAVIGTICFF RFWLGGGICG
 151 DYPLSATIMS EYSNKKTRCA FIAAVFAMQG VGIIFAGLVS MIVSGILLHY
 201 HPAPANKENH DRSWQDQMPA ADYMMRIYIM IGAFPALATF YRMKMPETA

1 MALPGNGIYV VGEAMTILMT AMRGSTRWNA TAYVDENDS LKLPLDLKH
 51 ELNRIEDLRQ IEPQVFLAPF LEVIRTADAT GPILSLALAS VNKLSYGLI
 101 DPTSPNLADI VERIADAVTH ARFMGTQSS DGVTFMVIE VJHTLIRSPB
 151 GAIVSNVSKC EYMLSCFKIS FEPRLSEILR RSAEKSILKM VLLFFMRLPQ

201 PAERSDTML QKRRTIGDAA SGATOEKLK KVAQAQAP RKSSAVEEP
 251 QTPSANLTV PGLKAPILA TTPASPANI LDMOKIQT PTITASGED
 301 ETVPEPVI QVESTSEPL LDGETGEATS TLAANSSEY INSVGFRTQ
 351 QSTDHVTSL SPYGLPEIQE LFRFLIICN PLDKONSDSM MHTELSLTV
 401 AFEVAADNG KYEGLELEK DDLCLNLISL LSSELSLIFA ADIQLCLLF
 451 ESLRHLKFO LEAYLRKLE IASDNPKTP YEMRELAIDN LLOIMRIPGF
 501 VTELYINDC DLYCTDMFES LTNLSKYTL SATNAVSTH IISMDTLSTV
 551 IDSLERNOCA SKNSNNRES LPEAPATG SRSHHNSGL EGIYIDGNS
 601 VAAEKVENI ASFINASSHR LRLQSGEGV GITSEQLAKV KOKRRLISOG
 651 TERNQREK GIQYLQEHGI LNAELDPQV ALFLRENPL DKMIGEYIS
 701 KKKVDSKIL INFVDSFTT GLRVQALRL YLEFRLGE APLIFVLEH
 751 FSDHMKQO DPEANDAF RLAYALIMN MDHNSNAKR LNPMTLED
 801 TKNRLGNGG EDFQEMLAQ VENAINEEI VMPAQTGLV RENTQMKVL
 851 RRGDTHDSH HYVHDASYD EIFNIWQAS LSALSFMDC STENGVRTL
 901 AGEKSAAIS AHYNLHSDP ALVTLCKFT TLLSVEGHE PAPANNETQ
 951 AVNGLNGKA QAAMRTVELL VHDYGDCLRE SWKHILLYL QLFRLKLPLK
 1001 SLIEVEDECE ANGKAMILE KPREKQESGL FSLYSFIS EGOEPTVEE
 1051 QDFTLGRKC IKECOLDOML QSKFYQLES LOELIKCYLA LKAPQGHKS
 1101 IGLPYAEDQT VFMMELVKI VVHNRDMIP LMPAVRDQY LLLMGASCG
 1151 YDYLNRCIV AVLKAIYLM RNBELCPIVL QSIKMLMLK PALLRLRSKQ
 1201 ISITIELLK TSNQNHSEQ DMOIIFNLLE CVGAGAVPPN YDDAQLPLP
 1251 NSAKSDGAI SGEEDATVP ERGYTSDSEI TKASAPAVS SPSAEMILV
 1301 NNKSELTTA SRQSPPSLS APPVNTLVN COLDHAFPA LFKWDSLAF
 1351 ITRSVAHITP YNFEACVRCI RIFVEACRDG GTRORRKLES AAKOKSSKR
 1401 SERRPGMASS ASSSNLTLLT GDPDNOING NAAEQEDLAQ RYEQLSIQL
 1451 DLMTLVTRT AQIFRMMAE GCTVPOSAL WSPGCPILQ GIARLMDRR
 1501 REVTHAISC LQORALIVHD LQTLSTGEMC SCFHQVLEPL LNELPESNA
 1551 AGOLDAALLE ESRIRATIM SKVFLQHLT LIEIGNAFNE LMLDILDYE
 1601 RFMKVGSPTL SEQMOELKN MLVMSVAV FHNQDGLSQ ALMELTWRI
 1651 GEFPLNKEE LHFHEDSITP AVSLDTQIR FAHSTGTTDY SLSLPLGYS
 1701 KRISISASK CDTLPATQK KPMAKLRSL KLLKAAVAA APOQOLPAVT
 1751 ILPROTVSN ELVVSATPP AATPLIGSPV ESPRSILQ PPMADVLOP
 1801 PSFVPAQPI VPPOPAVTD PTPSTLLPD LVNEATAAV QATTSTHS
 1851 POEAQOPASI VOQNTIVTN NTYNSYAIEV PMAPEETAQ FGQOQOOLLY
 1901 QQYVOQOAO QOOLPAPASD PAIWPISHL LAGNAVPSLP KMPQASIVHS
 1951 PAPYESQAA TSGAGTAAAD IYQEIYQNPY NLTLOQHPOQ QLMQOQOQOQ

2001 QOATGMANAF PAVATPANYF NVNVDPSIP PGSELLYGQO

11AA_SEQUENCE 1.0

ID ABB69566 standard; Protein: 337 AA.

AC ABB69566;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 35490.

KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

OS Drosophila melanogaster.

FN W0200171042-A2.

PD 27-SEP-2001.

PE 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL13669.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

PS Disclosure; SEQ ID NO 35490; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 337 AA;

ABB69566 Length: 337 May 30, 2002 09:25 Type: P Check: 7480 ..

1 MCRISEFFKD ISIGSVFTL TALSGERYCA IVPPLRIQT KPLVFTAVM

51 IMILAILLGM PSVLFSDIKS YVFTATGMM TIEVSPFRD PEYAKENAVAG

101 KALYYVLLPL SIIGALYIMM AKRLHMSARN MPEQOSQOS RTOARARLHV

151 ARMYVAFYVV FFICFFPYHV FELWYHFYPT AEEDEDEFWN VLRIVGFCTS

201 FLNSGVNPVA LYCVSGVFRQ HFNRYLCGIC VKRQPHLRQH STATGMMDNT

251 SVMSKRSTY VCGTAGNLRA SLHRNSNGV GAAGGCVGG VSGRGVSFH

301 RODSKPLOHG NAHGGAGGG SSGLAGAGRT AAVSEKR

11AA_SEQUENCE 1.0

ID AAU45398 standard; Protein: 89 AA.

AC AAU45398;

XX

DT 27-FEB-2002 (first entry)
 DE Propionibacterium acnes immunogenic protein #6294.
 XX
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KM dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 PN WO200181581-A2.
 XX
 XX 01-NOV-2001.
 PD
 XX 20-APR-2001; 2001WO-US12865.
 PF
 XX 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59525.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID NO 6593; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 89 AA;
 AU45398 Length: 89 May 30, 2002 09:25 Type: P Check: 183 ..
 1 NESVPRNYHT SGYIRGSRKA TSLKPSTWST AKASIHYSYP TWMNRRLNLC
 51 RSLGGRGASP SSSSORVSPS AFRHMDTLH EIRAGSYSP
 11AA_SEQUENCE 1.0
 ID ABG03782 standard; Protein; 146 AA.
 XX
 AC ABG03782;
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #3773.
 XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 XX Homo sapiens.
 OS
 XX WO200175067-A2.
 PN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US08631.
 PF
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PR
 XX (HYSE-) HYSEQ INC.
 PA
 PI Drmanac RT, Liu C, Tang YT;
 PI
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS67969.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID NO 34141; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 146 AA;
 ABG03782 Length: 146 May 30, 2002 09:25 Type: P Check: 3344 ..
 1 LKKIDRLFLK DFWQRFDSL KALAVKQORT VYRLTLKAW NDELOAYAO
 51 LVSLGNPDFI EVFKKIGEM WTWIDYNRFQ ELIQEYDSG GSKTFESAKDY
 101 MARTHFWALF GASERGFDPK DTRHSEERNK FKGLIFWFE IYLIFF
 11AA_SEQUENCE 1.0
 ID ABG03783 standard; Protein; 244 AA.
 XX
 AC ABG03783;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #3774.
 XX
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 OS

XX WO200175067-A2.
 PN 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 DR N-PSDB; AAS67970.
 DR
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 20; SEQ ID NO 34142; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AbG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SO Sequence 244 AA;
 ABG03783 Length: 244 May 30, 2002 09:25 Type: P Check: 8253 ..
 1 GIAGFVNIHL DSLSFLTGPV GYKARFEKX MTAKHCALSL VGEPIWYPEI
 51 NRFLKLIHQC KISSFLVTNA GPFAEIRNLE PVTOLYVRVD ASRKDSLAKKI
 101 DRPLFKDFWQ RFLDSYKALA VKYLORIGSR TPTMDTKIYSY CPAVHPAAPT
 151 DMKSWPSLFE VPTSLYCPF YLQLVESADA EGTOKYRRLT AYYIAPVTEP
 201 PLIRKEPSCL WKQAEFGDLG KHWMLVEQFS STRVQEHGVG. WGYE
 11AA_SEQUENCE 1.0
 ID ABG05207 standard; Protein: 636 AA.
 XX
 AC ABG05207;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #5198.
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.

XX WO200175067-A2.
 PN 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 DR N-PSDB; AAS69394.
 DR
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 20; SEQ ID NO 35566; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SO Sequence 636 AA;
 ABG05207 Length: 636 May 30, 2002 09:25 Type: P Check: 5356 ..
 1 MLTSYPTISOE VLEVDLQCE PRPLQALKRV SPLQNDGVKR GAENTDMLLI
 51 SEQDFRIKQ VSTMRSSESA SASKALICHG TYTLRMGCMC MYLSVSSSL
 101 ENPRGNSKRV VDSGDSNHY SEGREPSSLI RGALISRCRS KHLLEPSSVF
 151 LPLAPCKQOS QEVRPILMSH LCGMKRSKGN PREGSHGEI DLKHYNQDLIL
 201 GKDLMPETRV TYHWGKGNDQ TPQGLQDTGS ELMLIPDPK RHSPSPKVG
 251 SYGGQVINGV LAQVRLTVAA VSDVVSLEQI INISPGTYA AIDVANAFFS
 301 IPVMAHLKQ FAFSWOGQOV TETVLPQRQW EINPTKIQSP STSVKFLGIQ
 351 WYGACQDIPS KYKEKLHLIA PPTTKKEAQR LVDLPGFRQ HIHVGVLAM
 401 PIYQVTRKAA SFEMGPEQEK ALQOVQAAVQ AALPLGYPDP ADPLVLEVDA
 451 VMSPOQAPIG ESQRRPLGFV SKALPPSAGN YSPERQLEF CYWALLETEH
 501 LTMGHQVTMQ PELPIMNWP SNQSSYKVGH AQQSIITKKK WYTRDCARAH
 551 PEGTIRIHRN RNOGVEVEV PLITPSPDPL AKALLPVAPT LCSAGTGVLV

601 PEGGTLPDGD TTMIPLMKVL RLPPEHGLL LPLSQD

11AA_SEQUENCE 1.0
ID ABG05563 standard; Protein; 636 AA.
XX
AC ABG05563;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #5554.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
PI
PI
PI
DR WPI: 2001-639362/73.
DR N-PSDB; AAS69750.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 35922; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 636 AA:
ABG05563 Length: 636 May 30, 2002 09:25 Type: P Check: 5356 ..

1 MLSTYPSIQE VLEVLDLCOE PRPIQALKFV SPLQNDQVVR GAENDMDLWLI
51 SEQSDRIKIQ VSTMSESLA SAKALICHG TYTLRMGCWC MYLISVSPSL
101 ENRGSKSVY VDSGDSNHYI SEGREPSLI RGALISRCRS KHLLFFSYVF
151 LPLGPRCKQS QEVRLIMSH LCGMKSKKGN PRGGSHGEI DLHKYNQLLI

201 GKDLMPETRV TVHWKGNDO TFOGLQDRTGS ELMILPDPK RHGSPPVKG
251 SYGGQVINGV LAQVRLTVAA VSDVSLLEQ INISPGTWYA AIDVANAFPS
301 IPIYMAHLKQ PAFSWOGOOY TFTVLPQRGW EINTPTIQCP STSVKRLGIG
351 WYACODDIPS KYEKLHLILA PPTTKREAOR LVYDLEGFRQ HIPHGVLLM
401 PIYQVTRKAA SEFWGPEDEK ALQVQAAYQ AALPLGFYDP ADPLVLEVDA
451 VMSPOQAPIC ESQRRPLGFW SKALPESAGN YSPFEROLFV CYNALLETEH
501 LTMGHQVYTMQ PELIPMNWVP SNQSSYKVGH AQOHSIIKWK WIYRDCARAH
551 PEGTIRIHRN RNQGYEVEYV PLTIIPSDPL AKALLPYPAT LCSAGTGVLY
601 PEGGTLPDGD TTMIPLMKVL RLPPEHGLL LPLSQD

11AA_SEQUENCE 1.0
ID ABG06670 standard; Protein; 236 AA.
XX
AC ABG06670;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #6661.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
PI
PI
PI
DR WPI: 2001-639362/73.
DR N-PSDB; AAS70857.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 37029; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 236 AA;

ABG0670 Length: 236 May 30, 2002 09:25 Type: P Check: 936 ..

1 LSTLGAKRRK EPMSLRELRL GSSPGGCDL LEGLVMTLTF SSFAPRLRSP
51 VPASEAACSA PGLAAMLRRA SACAGTWSGR SRCSSHNTDV HTVASLIKLY
101 LRELPEPIV YAKYEDFLSC AKLSKEERA GYKELAKQVK SLPYVNTNLL
151 KYICRFIDEV QSYSGVKKMS YONLATVFGP NILRKVEDP LTIMEDLIKI
201 KEDKIERKGC EKEHRNDLND FKSLSLITFV AFYPLD

11AA_SEQUENCE 1.0
ID ABG06767 standard; Protein: 636 AA.

XX AC ABG06767;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #6758.

XX KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI: 2001-639362/73.

XX DR N-PSDB; AAS70934.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -

XX PS Claim 20; SEQ ID NO 37126; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 636 AA;

ABG06767 Length: 636 May 30, 2002 09:25 Type: P Check: 5356 ..

1 MLSTPIQGE VLEVLDICGE PRPLQALKEV SPLQNDGYKR GAETDMLLI
51 SEQSDFKIQK VSTMRSSEA SASKALICHG TYTLMGCMC MYLSVSPSL
101 ENPKGNSKVV VDSGDNHIV SEGEKPSILI RGALISRCRS KHLILFSVSF
151 LPLGPKCKQS QEVPRILMSH LCGMKSKKGN PREGSHBEI DLMRYNDLIL
201 GKDLMPETRV TVHWKGNDQ TEOGLQDTGS ELMILPGDPK RHCSPPVKVG
251 SYGGGVINGV LAQVRLTVAA VSDVVSLLBQ INISPGTWYA AIDVANAFS
301 IPVHAHLKQ FAFSWGQQT TTVLPPQGW EINPKTIQGP STSVKFLGIQ
351 WYGACQDIPS KYKELHLA PPTTKREAQR LVDLGFFRRQ HIPHVGVLW
401 PIYQVTRKAA SFEMGPEQEK ALQOVQAAVQ AALPLGYPDP ADPLVLEVDA
451 VWSPOQAPIG ESQRRPLGFV SKALPPSAGN YSPFEROLF CYMALLETEH
501 LTMGHQVTM PELPIMMWP SNOSSYKVGH AOOHSIIKWK WYIRDCARAH
551 PEGTIRIIRS RNOGVEVEV PLITPSDPL AKALLPVPAT LCSAGTGVLV
601 PEGTLPPEGD TTKLPLMWKL RLPFGHGLL LPLSOD

11AA_SEQUENCE 1.0
ID ABG17402 standard; Protein: 636 AA.

XX AC ABG17402;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #17393.

XX KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI: 2001-639362/73.

XX DR N-PSDB; AAS81589.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -

XX PS Claim 20; SEQ ID NO 47761; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 636 AA:

ABG17402 Length: 636 May 30, 2002 09:25 Type: P Check: 5356 ..

1 MLSTYISQD VLEVDLCOE PRPIQALKFY SPLQNDGVK RAEVDLMLI
 51 SEQSDFKIQR VSTMSESTIA SASKALICHG TYTLRMCCMC MYLSVSPSSL
 101 ENRGNKSKVY VDSGNSHIY SEKKEPSSLI RGALISRCRS KHLLEFSYF
 151 LPLGPKCKOS QEVRLIMSH LCGWKSCKGN PREGSGEIT DLHKYQQLL
 201 GNDLMPFTRY TVHWKGNQD TFOGIDTGS ELMILPDPK RHCSPVKKV
 251 SYGGQVINGV LAQVRLTVAA VSDVVSLEQ INISPGTWYA AIDVANAFPS
 301 IPYHMAHLKQ FAFSWGGQY TETVLPQGW EIMPTKIQG STSVKRLGIO
 351 WYACQDIIIS KYEKLHLILA PPTTKKEAQR LVDLFGFRQ HIPHGVLLM
 401 PIYQVTRKKA SFEMGPEDEK ALQOVQAAVQ AALPLGPYP ADPLVEVDA
 451 VMSPOQAPIG ESQRRPLGFV SKALPSSAGN YSPFERQLF CYWALLETEH
 501 LTMGHQVYMO PELPTMNVV SNOSYKVGH AOOHSIIKKW WYRDCARAH
 551 PECTIRIHRN RNQGVVEVY PLTTPSDPL AKALLPYPAT LCSAGTGVLY
 601 PEGGLPQGD TTMIPLNWKL RLPGHGGL LPLSQD

IIAA_SEQUENCE 1.0
 ID ABB28764 standard; Protein; 198 AA.

XX AC ABG29764;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #29755.
 XX KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217;
 PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YF;
 XX WPI: 2001-639362/73.
 DR N-PSDB: AAS93951.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

PS Claim 20. SEQ ID No 60123; 103pp: English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 198 AA:

ABG29764 Length: 198 May 30, 2002 09:25 Type: P Check: 9364 ..

1 VRAVFLSREQ IRSHEIRISA AVSNHQHRR ACHIDRRTV QTLAHLTFRF
 51 GDKGVTRPED FVHFHFRRT KQDPLTFWP SAPVTATILR IPLANGFFRH
 101 DHRSSCMTRV PQVEDPVKS HAPQDFWTH GKMFAYDVA IGIPFPAGR
 151 SCVILLGDAV REVTRVSNV PAKRYIGSP RRGFINSCNM PGIIISGA

IIAA_SEQUENCE 1.0
 ID ABB29544 standard; Peptide; 27 AA.

XX AC ABB29544;
 XX DT 01-FEB-2002 (first entry)
 XX DE Peptide #2195 encoded by breast cell single exon nucleic acid probe.
 XX KM Human: microarray; single exon probe; gene expression; breast;
 XX KW disease; cancer.
 XX OS Homo sapiens.
 XX PN WO200157271-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00662.
 XX PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-496933/54.
 XX
 PT New spatially-addressable set of single exon nucleic acid probes,
 PT useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes -
 PS
 XX Claim 27; SEQ ID NO 12512; 327pp + sequence listing; English.
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting
 CC the probes with a collection of detectably labelled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to
 CC encode proteins. They are useful for gene discovery, and for
 CC determining predisposition and/or prognosing breast disease. Gene
 CC expression analysis is useful for assessing the toxicity of chemical
 CC agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression, with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a peptide encoded by a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 27 AA;
 ABB29544 Length: 27 May 30, 2002 09:25 Type: P Check: 8372 ..
 1 SLKKIDRPLF KDFMQFLDS LKALAVK
 IIAA_SEQUENCE 1.0
 ID ABB34423 standard; Peptide; 113 AA.
 AC ABB34423;
 XX
 DT 04-FEB-2002 (first entry)
 XX
 DE Peptide #1929 encoded by human foetal liver single exon probe.
 XX
 KM Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 PD
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX

DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 PS
 XX Claim 27; SEQ ID NO 27058; 639pp + sequence listing; English.
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 113 AA;
 ABB34423 Length: 113 May 30, 2002 09:25 Type: P Check: 1841 ..
 1 FITSLPDM EATLWIGLRMT AYEKINKMTD NRELTYSNF PLVSGRLRI
 51 PENVSFFRNQ YSSYLCEILE SEIHPTKON PTVLFHLYEM SRIGTSKIYS
 101 RLVIARGYGG MKD
 IIAA_SEQUENCE 1.0
 ID ABB19835 standard; Protein; 113 AA.
 AC ABB19835;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Protein #1834 encoded by probe for measuring heart cell gene expression.
 XX
 KM Human; gene expression; heart; microarray; vascular system;
 KM cardiovascular disease; hypertension; cardiac arrhythmia;
 KM congenital heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157274-A2.
 PD
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00666.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488899/53.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 PS
 XX Claim 15; SEQ ID NO 21605; 530pp; English.
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,

CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 113 AA:

ABBI9835 Length: 113 May 30, 2002 09:25 Type: P Check: 1841 ..

1 FITSLPDM EATLMIGLRMT AYERKINKWTD NRELYTSNFH PLLVSGRLRT

51 PEVSEFFKND YSSYLCEILLE SFHFHTKON PTVLFHLYEM SKIGSKIYS

101 RLVIANGYGG WKD

IIAA_SEQUENCE 1.0

ID AAB86968 standard; Protein: 506 AA.

AC AAB86968;

DT 11-DEC-2001 (first entry)

DE D. melanogaster peptide receptor SEQ ID 34.

KM Insect; fruitfly; peptide receptor; plant protection; insecticide.

OS Drosophila melanogaster.

PN DE10013618-A1.

PD 20-SEP-2001.

PF 18-MAR-2000; 2000DE-1013618.

PR 18-MAR-2000; 2000DE-1013618.

PA (FARB) BAYER AG.

PI Antonicek H, Friedrich G, Schulte T;

DR MPI: 2001-571695/65.

DR N-PSDB: AAH49421.

PT New polypeptides from Drosophila melanogaster have biological activity
 PT of peptide receptor, useful to find new compounds for plant protection
 PT and insecticides -

PS Claim 2: Page 90-92; 128pp; German.

CC This invention describes novel polypeptides (PI) from Drosophila
 CC melanogaster having the biological activity of a peptide receptor.
 CC Molecules of the invention are used to find new plant protection
 CC compounds or insecticides, or to find genes encoding a polypeptide
 CC involved in the structure of functionally similar receptors in insects
 CC This sequence represents a Drosophila melanogaster (fruitfly) peptide
 CC receptor described in the method of the invention.

XX Sequence 506 AA:

AAB86968 Length: 506 May 30, 2002 09:25 Type: P Check: 4326 ..

1 WIASLMDVQO TLARLADSD GNGANDSGL ATGQGLEQDO ECIALDMGN

51 ASADGIVPY VPVLDREPTY IVTVLYTLIF YGVLGNGTL VIFFRHRS

101 RNIPNYIIS LALADLVIL VCVVPAIVY TQSNPFERN MGRISEFFRD

151 ISIGSVFTL TALSGERYCA IVNPLRLQOT KPIVFTAVM IMIILAILGM

201 PSYLFSDIKS YPVFTATGNN TIEVCSPPRD PEYAKFMVAG KALVYVLLPL

251 SITGALITM AKRLHSAFN MPGEQSQMS RTQARARLHV ARNVAVFVV
 301 FFLCFEPYHV FELWYHFPPT AEEDEFDEFWN VLRIVGFCST AVLKNNVSA
 351 IMWLNSEPLV VKAICILAKW LPKLVROPRG LYCVSGVFRQ HFNRYLCCIC
 401 VKRQPHLRQH STRATGMNDT SVMSKRIRSTY VGTAGNLRA SLHRNSNHCV
 451 GGAGGAGVGG VSGRVSPFH RQDSMPLOHG NAROGAGAGG SSGLAGAGRT
 501 AAVSEK

IIAA_SEQUENCE 1.0

ID AAE11926 standard; Protein: 288 AA.

AC AAE11926;

DT 18-DEC-2001 (first entry)

DE Human CG27 (or C869) receptor protein #1.

KM Human; apolipoprotein; lipase; lipoprotein receptor; ALR; angina;
 KM cardiovascular disease; lipid metabolism; myocardial infarction;
 KM cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
 KM coronary artery thrombosis; cerebral artery thrombosis; stroke;
 KM intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
 KM neuroprotectant; cerebroprotective.

XX Homo sapiens.

OS Homo sapiens.

PN WO200179446-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US12529.

PR 14-APR-2000; 2000US-197337P.

PR 20-JUN-2000; 2000US-0598042.

PR 03-AUG-2000; 2000US-0631451.

PR 22-SEP-2000; 2000US-0667298.

PR 17-NOV-2000; 2000US-0714936.

PA (HYSE-) HYSEQ INC.

PI Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
 PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
 PI Wang D;

DR MPI: 2001-611724/70.

DR N-PSDB: AAD19221.

PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
 PT receptor polypeptides, useful for preventing diagnosing and treating
 PT lipid metabolism disorders, thrombosis and cardiovascular diseases -

PS Claim 10: Fig 4; 266pp; English.

CC The invention relates to polynucleotides encoding proteins CG122, CG179,
 CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
 CC involved in lipid metabolism and cardiovascular disease such as human
 CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
 CC and protein sequences are useful for treating or preventing disorders
 CC associated with apolipoproteins, lipases and lipoprotein receptor (ALR)
 CC expression and for treating lipid metabolism, cardiovascular diseases
 CC and thrombosis. Antibodies against these proteins are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of these sequences. ALR polypeptides are also
 CC useful for identifying agents (agonists and antagonists) that bind to
 CC them and cells expressing ALR proteins are useful for identifying a
 CC therapeutic agent for use in treatment of a pathology related to
 CC aberrant expression or physiological interactions of this polypeptide.
 CC Vectors comprising these DNA and protein sequences are also useful for
 CC producing ALR proteins. The nucleic acids and polypeptides of the

CC Invention are also useful for the treatment of occlusive cardiovascular
 CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
 CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
 CC or intracardiac thrombosis and stroke. The nucleotides of the invention
 CC are used in gene therapy. The present sequence is human CG27 (or C869)
 CC receptor protein.

XX Sequence 288 AA;

AAE11926 Length: 288 May 30, 2002 09:25 Type: P Check: 2953 ..

1 MSLHSQASAT TRHPEPRRTE HRAPSTWRP VALTLTLCL VLLIGLALG
 51 LLEFQYVOLS NTGODITSQE EERLGTSGE LQSLQVONIK LAGSLQHYAE
 101 KLCELYNKA GGYTRNNVPA SASSESLROL PHMGSAANAH RCPCTEQWK
 151 WHGDNCTQIFY KDSKSWEDCK YFCLSENSTM LKINKQEDLE FASQSYSEF
 201 FYSYWTGLLR PDSGKAWLMV DGTPTSELE HIIIDVTSPR SRDCVAILNG
 251 MIFSNDCKEL KRCYCERRAG MKRPESLHVP PETLGECD

!!AA_SEQUENCE 1.0
 ID AAE11932 standard; Protein; 280 AA.

AC AAE11932;

DT 18-DEC-2001 (first entry)

DE Human CG27 (or C869) receptor protein variant #2.

KM Human; apolipoprotein; lipase; lipoprotein receptor; ALAr; angina;
 KM cardiovascular disease; lipid metabolism; myocardial infarction;
 KM cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
 KM coronary artery thrombosis; cerebral artery thrombosis; stroke;
 KM intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
 KM neuroprotectant; cerebroprotective; variant.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Domain 148..166

PN /note="C-type lectin domain"

MO200179446-A2.

PD 25-OCT-2001.

PE 16-APR-2001; 2001WO-US12529.

PR 14-APR-2000; 2000US-197137P.

PR 20-JUN-2000; 2000US-0598042.

PR 03-AUG-2000; 2000US-0631451.

PR 22-SEP-2000; 2000US-0667298.

PR 17-NOV-2000; 2000US-0714936.

PA (HYSE-) HYSEQ INC.

PI Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;

PI Lin C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;

PI Wang D;

DR MPI: 2001-611724/70.

DR N-PSDB: AAD19227.

PF Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein

PF receptor polypeptides, useful for preventing diagnosing and treating

PF lipid metabolism disorders, thrombosis and cardiovascular diseases -

XX Claim 10; Fig 4; 266pp; English.

XX The invention relates to polynucleotides encoding proteins CG122, CG179,

CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
 CC involved in lipid metabolism and cardiovascular disease such as human
 CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
 CC and protein sequences are useful for treating or preventing disorders
 CC associated with apolipoproteins, lipases and lipoprotein receptor (ALAr)
 CC expression and for treating lipid metabolism, cardiovascular diseases
 CC and thrombosis. Antibodies against these proteins are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of these sequences. ALAr polypeptides are also
 CC useful for identifying agents (agonists and antagonists) that bind to
 CC them and cells expressing ALAr proteins are useful for identifying a
 CC therapeutic agent for use in treatment of a pathology related to
 CC aberrant expression or physiological interactions of this polypeptide.
 CC Vectors comprising these DNA and protein sequences are also useful for
 CC producing ALAr proteins. The nucleic acids and polypeptides of the
 CC invention are also useful for the treatment of occlusive cardiovascular
 CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
 CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
 CC or intracardiac thrombosis and stroke. The nucleotides of the invention
 CC are used in gene therapy. The present sequence is human CG27 (or C869)
 CC receptor protein variant.

XX Sequence 280 AA;

AAE11932 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..

1 MQAKYSTFRD MLDDGDPTM SLHSQASATP RHPEPRRTEH RAPSTWRPV
 51 ALTTLTLCLV LLIGLALGL LFFQYVOLS NTGODITSQE ERLGNTSGEL
 101 OSLOVONIKL AGSLQHYAEK LCRELYNKAAG AHRCSPCTBQ WKMHQDNCYQ
 151 FKDSKSWED CKYFCLSENS TMLKINKQED LEFASQSYVS EEFYSYWTGL
 201 LRPDSGKAWL WMDGTPTFSE LFHIIIDVTS PMSRDCVAIL NGMIFSKDCK
 251 ELKRCYCERR AGWKPESLH VPPTLGECD

!!AA_SEQUENCE 1.0
 ID AAE11933 standard; Protein; 314 AA.

AC AAE11933;

DT 18-DEC-2001 (first entry)

DE Human CG27 (or C869) receptor protein variant #3.

KM Human; apolipoprotein; lipase; lipoprotein receptor; ALAr; angina;
 KM cardiovascular disease; lipid metabolism; myocardial infarction;
 KM cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
 KM coronary artery thrombosis; cerebral artery thrombosis; stroke;
 KM intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
 KM neuroprotectant; cerebroprotective; variant.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Domain 190..284

PN /note="C-type lectin domain"

MO200179446-A2.

PD 25-OCT-2001.

PE 16-APR-2001; 2001WO-US12529.

PR 14-APR-2000; 2000US-197137P.

PR 20-JUN-2000; 2000US-0598042.

PR 03-AUG-2000; 2000US-0631451.

PR 22-SEP-2000; 2000US-0667298.

PR 17-NOV-2000; 2000US-0714936.

PA (HYSE-) HYSEQ INC.

XX Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R,
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;
XX WPI: 2001-611724/70.
DR N-PSDB: AAD19228.
XX
PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX
PS Claim 10; Fig 4; 266pp; English.
XX
XX The invention relates to polynucleotides encoding proteins CG122, CG179,
CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
CC involved in lipid metabolism and cardiovascular disease such as human
CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
CC and protein sequences are useful for treating or preventing disorders
CC associated with apolipoproteins, lipases and lipoprotein receptor (ALR)
CC expression and for treating lipid metabolism, cardiovascular diseases
CC and thrombosis. Antibodies against these proteins are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of these sequences. ALR polypeptides are also
CC useful for identifying agents (agonists and antagonists) that bind to
CC them and cells expressing ALR proteins are useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to
CC aberrant expression or physiological interactions of this polypeptide.
CC Vectors comprising these DNA and protein sequences are also useful for
CC producing ALR proteins. The nucleic acids and polypeptides of the
CC invention are also useful for the treatment of occlusive cardiovascular
CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
CC or intracardiac thrombosis and stroke. The nucleotides of the invention
CC are used in gene therapy. The present sequence is human CG27 (or C869)
CC receptor protein variant.
XX
SQ Sequence 314 AA:
AAE11933 Length: 314 May 30, 2002 09:25 Type: P Check: 6095 ..
1 MQAKYSTMD MLDDEDTM SLHSQASAT RHPEPRTEH RAPSITRPV
51 ALTLTLCLV LLGLAALGL LFFQYQSLN TGQDTISGM ERLGNTQEL
101 QSLQVQNIKL AGSLQVAAEK LCRLYNKAG GYTRNWPAS ASSESIRQLP
151 HMGESAAHR CSPCTEQMKW HGDNCYQFYK DSKSWEDCKY FCLSENSTML
201 KINKQEDLEF AASQSYSEFF YSYWTGLLRP DSGKAWLMMD GTPPTSLEFH
251 IIDVTSPRS RDCVAILNGM IFSKDKELK RVCERRAGM VKPESLHVP
301 ETLEGDMHH HHHH
ID AAE11934 standard; Protein: 247 AA.
XX
AC AAE11934;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human CG27 (or C869) receptor protein variant #4.
XX
XX Human; apolipoprotein; lipase; lipoprotein receptor; ALR; angina;
KW cardiovascular disease; lipid metabolism; myocardial infarction;
KW cerebral ischaemia; arterial thrombosis; thrombotic; antilipemic;
KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
KW neuroprotectant; cerebroprotective; variant.
XX
OS Homo sapiens.
XX

EH Key
FT Domain
FT Location/Qualifiers
FT 130..224
FT /note="C-type lectin domain"
XX
XX WO200179446-A2.
XX
XX 25-OCT-2001.
XX
XX 16-APR-2001; 2001WO-US12529.
XX
XX 14-APR-2000; 2000US-197137P.
XX 20-JUN-2000; 2000US-059804Z.
XX 03-AUG-2000; 2000US-0631451.
XX 22-SEP-2000; 2000US-0667298.
XX 17-NOV-2000; 2000US-0714936.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R,
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;
XX WPI: 2001-611724/70.
DR N-PSDB: AAD19229.
XX
XX Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX
PS Claim 10; Fig 4; 266pp; English.
XX
XX The invention relates to polynucleotides encoding proteins CG122, CG179,
CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
CC involved in lipid metabolism and cardiovascular disease such as human
CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
CC and protein sequences are useful for treating or preventing disorders
CC associated with apolipoproteins, lipases and lipoprotein receptor (ALR)
CC expression and for treating lipid metabolism, cardiovascular diseases
CC and thrombosis. Antibodies against these proteins are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of these sequences. ALR polypeptides are also
CC useful for identifying agents (agonists and antagonists) that bind to
CC them and cells expressing ALR proteins are useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to
CC aberrant expression or physiological interactions of this polypeptide.
CC Vectors comprising these DNA and protein sequences are also useful for
CC producing ALR proteins. The nucleic acids and polypeptides of the
CC invention are also useful for the treatment of occlusive cardiovascular
CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
CC or intracardiac thrombosis and stroke. The nucleotides of the invention
CC are used in gene therapy. The present sequence is human CG27 (or C869)
CC receptor protein variant.
XX
SQ Sequence 247 AA:
AAE11934 Length: 247 May 30, 2002 09:25 Type: P Check: 89 ..
1 MQAKYSTMD MLDDEDTM SLHSQASAT RHPEPRTEH RAPSITRPV
51 DTISQMEERL GNTSQELSL QYQNTIKLAGS LQVAAEKLR ELVKNKAGHR
101 CSPCTEQMKW HGDNCYQFYK DSKSWEDCKY FCLSENSTML KINKQEDLEF
151 AASQSYSEFF YSYWTGLLRP DSGKAWLMMD GTPPTSLEFH IIDVTSPRS
201 RDCVAILNGM IFSKDKELK RVCERRAGM VKPESLHVP ETLEGDM
ID AAE11938 standard; Protein: 280 AA.
XX
XX AAE11938;
XX

DT 18-DEC-2001 (first entry)
 XX Human lipid metabolism related protein #1.
 XX
 KM Human: apolipoprotein; lipase; lipoprotein receptor; ALR; angina;
 KM cardiovascular disease; lipid metabolism; myocardial infarction;
 KM cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
 KM coronary artery thrombosis; cerebral artery thrombosis; stroke;
 KM intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
 KM neuroprotectant; cerebroprotective.
 XX
 OS Homo sapiens.
 XX
 PN MO200179446-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001MO-US12529.
 XX
 PR 14-APR-2000; 2000US-197137P.
 PR 20-JUN-2000; 2000US-0598042.
 PR 03-AUG-2000; 2000US-0631451.
 PR 22-SEP-2000; 2000US-0667298.
 PR 17-NOV-2000; 2000US-0714936.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
 PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
 PI Wang D;
 DR WPI; 2001-611724/70.
 DR N-PSDB; AAD19233.
 XX
 PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
 PT receptor polypeptides, useful for preventing diagnosing and treating
 PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
 XX
 PS Claim 10; Page 252; 266pp; English.
 XX
 CC The invention relates to polynucleotides encoding proteins CGI22, CGI179,
 CC CG95, CGI21, CGI62, CG27, CGI53 and CGI68 which are related to proteins
 CC involved in lipid metabolism and cardiovascular disease such as human
 CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
 CC and protein sequences are useful for treating or preventing disorders
 CC associated with apolipoproteins, lipases and lipoprotein receptor (ALR)
 CC expression and for treating lipid metabolism, cardiovascular diseases
 CC and thrombosis. Antibodies against these proteins are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of these sequences. ALR polypeptides are also
 CC useful for identifying agents (agonists and antagonists) that bind to
 CC them and cells expressing ALR proteins are useful for identifying a
 CC therapeutic agent for use in treatment of a pathology related to
 CC aberrant expression or physiological interactions of this polypeptide.
 CC Vectors comprising these DNA and protein sequences are also useful for
 CC producing ALR proteins. The nucleic acids and polypeptides of the
 CC invention are also useful for the treatment of occlusive cardiovascular
 CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
 CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
 CC or intracardiac thrombosis and stroke. The nucleotides of the invention
 CC are used in gene therapy. The present sequence is human protein
 CC related to proteins involved in lipid metabolism.
 XX
 SO Sequence 280 AA;

AAE11938 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..

1 MQARYSTRD MLDDGDTM SLHSQASATT RHPEPRTEH RAPSTWRPV
 51 ALTLTLCTLV LILGIAIIGL LFGYYVQLSN TGQDITSGME ERLGNTSGEL
 101 QSLQVQNIKL AGSLQVHAEK LCRELYNKAG AHRSPCTEQ WKWHGDMCIQ

151 FYKSKSWED CKYFCUSENS TMKINKQED LEPAASQSY EPHYWTGL
 201 LRPDSGKAML WMDGTPETSE LPHIIDVYS PRSRDCAVAIL NGMIFSKCK
 251 ELKRCVCERR AGWVKPESLH VPPELTGEGD
 I1AA_SEQUENCE 1.0
 ID AAE11939 standard; Protein: 307 AA.
 XX
 AC AAE11939;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human lipid metabolism related protein #2.
 XX
 KM Human: apolipoprotein; lipase; lipoprotein receptor; ALR; angina;
 KM cardiovascular disease; lipid metabolism; myocardial infarction;
 KM cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
 KM coronary artery thrombosis; cerebral artery thrombosis; stroke;
 KM intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
 KM neuroprotectant; cerebroprotective.
 XX
 OS Homo sapiens.
 XX
 PN MO200179446-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001MO-US12529.
 XX
 PR 14-APR-2000; 2000US-197137P.
 PR 20-JUN-2000; 2000US-0598042.
 PR 03-AUG-2000; 2000US-0631451.
 PR 22-SEP-2000; 2000US-0667298.
 PR 17-NOV-2000; 2000US-0714936.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
 PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
 PI Wang D;
 DR WPI; 2001-611724/70.
 DR N-PSDB; AAD19234.
 XX
 PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
 PT receptor polypeptides, useful for preventing diagnosing and treating
 PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
 XX
 PS Claim 10; Page 254-255; 266pp; English.
 XX
 CC The invention relates to polynucleotides encoding proteins CGI22, CGI179,
 CC CG95, CGI21, CGI62, CG27, CGI53 and CGI68 which are related to proteins
 CC involved in lipid metabolism and cardiovascular disease such as human
 CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
 CC and protein sequences are useful for treating or preventing disorders
 CC associated with apolipoproteins, lipases and lipoprotein receptor (ALR)
 CC expression and for treating lipid metabolism, cardiovascular diseases
 CC and thrombosis. Antibodies against these proteins are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of these sequences. ALR polypeptides are also
 CC useful for identifying agents (agonists and antagonists) that bind to
 CC them and cells expressing ALR proteins are useful for identifying a
 CC therapeutic agent for use in treatment of a pathology related to
 CC aberrant expression or physiological interactions of this polypeptide.
 CC Vectors comprising these DNA and protein sequences are also useful for
 CC producing ALR proteins. The nucleic acids and polypeptides of the
 CC invention are also useful for the treatment of occlusive cardiovascular
 CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
 CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
 CC or intracardiac thrombosis and stroke. The nucleotides of the invention
 CC are used in gene therapy. The present sequence is human protein
 CC related to proteins involved in lipid metabolism.

XX Sequence 307 AA;
 S0
 AAE11939 Length: 307 May 30, 2002 09:25 Type: P Check: 2921 ..
 1 MQAKYSSTRD MDDDDGDTTM SLHSQASATP RHPEPRRTVE RASSTWRPV
 51 ALPLTLTCLV LILGIALGL LFPYQYQJLN TGQDTTISOME ERIGNTSQEL
 101 QSLQVONIKL AGSLQHVAKR LCRELNNKAG GYTRNNVPAS ASSESIRQLP
 151 HMGSAANAHR CSPCTEQWKW HGDNCYQFYK DSKSWEDOCK FCLSENSTML
 201 KINKQEDLEF AASQSYSEFF YSYMTGLLRP DSGKAMLMMD GTPPTSELPH
 251 IIDVTSPPRS RDCVALLNGM IFSKCKEIK RCVCERRAGM VKPESLHVPV
 301 ETLEGGD
 11AA_SEQUENCE 1.0
 ID AAE11943 standard; Protein: 281 AA.
 AC AAE11943;
 XX
 DT 18-DEC-2001 (first entry)
 DE Human CG27 (or C869) receptor protein variant #5.
 XX
 XX Human; apolipoprotein; lipase; lipoprotein receptor; ALIr; angina;
 KM cardiovascular disease; lipid metabolism; myocardial infarction;
 KM cerebral ischaemia; arterial thrombosis; thrombolytic; antiplatelet;
 KM coronary artery thrombosis; cerebral artery thrombosis; stroke;
 KM intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
 KM neuroprotectant; cerebroprotective; variant.
 XX
 OS Homo sapiens.
 XX
 PN WO200179446-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US12529.
 XX
 PR 14-APR-2000; 2000US-197137P.
 PR 20-JUN-2000; 2000US-0598042.
 PR 03-AUG-2000; 2000US-0631451.
 PR 22-SEP-2000; 2000US-0667298.
 PR 17-NOV-2000; 2000US-0714936.
 XX
 PA (HYSR-) HYSER INC.
 XX
 PI Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
 PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
 PI Wang D;
 XX
 DR WPI: 2001-611724/70.
 XX
 N-PSDB: AAD19238.
 XX
 PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
 PR receptor polypeptides, useful for preventing diagnosing and treating
 PR lipid metabolism disorders, thrombosis and cardiovascular diseases -
 XX
 PS Claim 10; Page 265-266; 266pp; English.
 XX
 XX The invention relates to polynucleotides encoding proteins CG122, CG179,
 CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
 CC involved in lipid metabolism and cardiovascular disease such as human
 CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
 CC and protein sequences are useful for treating or preventing disorders
 CC associated with apolipoproteins, lipases and lipoprotein receptor (ALIr)
 CC expression and for treating lipid metabolism, cardiovascular diseases
 CC and thrombosis. Antibodies against these proteins are useful for
 CC determining the presence of or predisposition to a disease associated

CC with altered levels of these sequences. ALIr polypeptides are also
 CC useful for identifying agents (agonists and antagonists) that bind to
 CC them and cells expressing ALIr proteins are useful for identifying a
 CC therapeutic agent for use in treatment of a pathology related to
 CC aberrant expression or physiological interactions of this polypeptide.
 CC Vectors comprising these DNA and protein sequences are also useful for
 CC producing ALIr proteins. The nucleic acids and polypeptides of the
 CC invention are also useful for the treatment of occlusive cardiovascular
 CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
 CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
 CC or intracardiac thrombosis and stroke. The nucleotides of the invention
 CC are used in gene therapy. The present sequence is human CG27 (or C869)
 CC receptor protein variant.
 XX
 S0 Sequence 281 AA;
 AAE11943 Length: 281 May 30, 2002 09:25 Type: P Check: 7726 ..
 1 MQAKYSSTRD MDDDDGDTTM SLHSQASATP RHPEPRRTVE QYQJLSNTGQ
 51 DTISQMEERL GNTSOELQSL QVQNIKIAGS LQHVAKLCR ELYNKAGGYT
 101 RNWVPASASS ESRLQPLHMG ESAAHRCSP CTEQMKHGD NCQFYKDSK
 151 SWEDCKTFCL SENSTMKLN KOEDLEFAAS QSYSEFFYSY WTGILRPDSG
 201 KAWLMMDGTP FTSELFHIII DVTSPSRDC VAILNGMFS KCKEKLRCV
 251 CERRAGWKP ESLHVPPELT GEGDMHHHH H
 11AA_SEQUENCE 1.0
 ID AAU20199 standard; Protein: 85 AA.
 XX
 AC AAU20199;
 XX
 DT 17-DEC-2001 (first entry)
 DE Human novel endocrine antigen, SEQ ID No 256.
 XX
 XX Human
 KM Human; endocrine antigen; cytostatic; antifertility; antidiabetic;
 KM thyroid-active; adrenal-active; androgenic; gastric; gene therapy;
 KM antisense-therapy; antibody; endocrine disorder; hormone imbalance;
 KM reproductive disorder; endocrine cancer; pancreatic disorder;
 KM diabetes mellitus; adrenal gland disorder; hirsutism; thyroid disorder;
 KM hyperthyroidism; hypohalamic disorder; vanishing testes syndrome.
 OS Homo sapiens.
 XX
 PN WO200155319-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01335.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.

PR	14-AUG-2000	200005-02234513
PR	14-AUG-2000	200005-02235113
PR	14-AUG-2000	200005-02252114
PR	14-AUG-2000	200005-02252566
PR	14-AUG-2000	200005-02253567
PR	14-AUG-2000	200005-02255168
PR	14-AUG-2000	200005-02255170
PR	14-AUG-2000	200005-02254470
PR	14-AUG-2000	200005-02255158
PR	14-AUG-2000	200005-02255158
PR	14-AUG-2000	200005-02257159
PR	18-AUG-2000	200005-02262759
PR	22-AUG-2000	200005-02266811
PR	22-AUG-2000	200005-02266868
PR	22-AUG-2000	200005-02271882
PR	30-AUG-2000	200005-02270024
PR	30-AUG-2000	200005-02270929
PR	01-SEP-2000	200005-02292887
PR	01-SEP-2000	200005-02293437
PR	01-SEP-2000	200005-02293438
PR	08-SEP-2000	200005-02311442
PR	08-SEP-2000	200005-02311443
PR	08-SEP-2000	200005-02311443
PR	08-SEP-2000	200005-02311444
PR	14-SEP-2000	200005-02323298
PR	14-SEP-2000	200005-02323299
PR	14-SEP-2000	200005-02324200
PR	14-SEP-2000	200005-02324201
PR	14-SEP-2000	200005-02330363
PR	14-SEP-2000	200005-02330364
PR	21-SEP-2000	200005-02343223
PR	21-SEP-2000	200005-02344774
PR	25-SEP-2000	200005-02349737
PR	25-SEP-2000	200005-02349898
PR	26-SEP-2000	200005-02354884
PR	27-SEP-2000	200005-02358534
PR	27-SEP-2000	200005-02358536
PR	29-SEP-2000	200005-02363127
PR	29-SEP-2000	200005-02363167
PR	29-SEP-2000	200005-02365658
PR	29-SEP-2000	200005-02365659
PR	02-OCT-2000	200005-02368670
PR	02-OCT-2000	200005-02370337
PR	02-OCT-2000	200005-02370338
PR	02-OCT-2000	200005-02370339
PR	02-OCT-2000	200005-02370349
PR	13-OCT-2000	200005-02393935
PR	13-OCT-2000	200005-02393937
PR	20-OCT-2000	200005-02409660
PR	20-OCT-2000	200005-02411821
PR	20-OCT-2000	200005-02411821
PR	20-OCT-2000	200005-02417866
PR	20-OCT-2000	200005-02417867
PR	20-OCT-2000	200005-02418087
PR	20-OCT-2000	200005-02418089
PR	20-OCT-2000	200005-02418206
PR	01-NOV-2000	200005-02446174
PR	08-NOV-2000	200005-02446174
PR	08-NOV-2000	200005-02446475
PR	08-NOV-2000	200005-02446475
PR	08-NOV-2000	200005-02446476
PR	08-NOV-2000	200005-02464776

PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249267.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.
XX		
XX	Rosen CA,	Barash SC, Ruben SM;
PI		
XX	WPI: 2001-457726/49.	
DR		
DR	N-PSDB: AAS32302.	
XX		
PT	Isolated polypeptide for treating, preventing and prognosing disorders	
PT	related to the endocrine system including endocrine disorders,	
PT	reproductive disorders, and gastrointestinal disorders and also for	
PT	testing and detection e.g. diagnosis -	
PI		
XX		
PS	Claim 11; SEQ ID NO 256; 558bp; English.	
XX		
CC	The invention relates to cDNAs encoding novel human endocrine	
CC	antigens or a fragment having biological activity, a domain, an epitope,	
CC	full length protein, variant, allelic variant or a species homologue of	
CC	the cDNA/antigen. The DNAs and polypeptides are useful for preventing,	
CC	treating or ameliorating a medical condition when administered	
CC	(e.g. by gene therapy or antisense-therapy). Identifying mutations in	
CC	the genes coding for the antigens is useful for diagnosing a pathological	
CC	condition or a susceptibility to a pathological condition. The DNAs,	
CC	antigens and antibodies raised against the antigens useful for treating,	
CC	preventing and/or prognosing disorders related to the endocrine system	
CC	or hormone imbalance or reproductive disorders, cancers of endocrine	
CC	tissues, disorders of the pancreas (e.g. diabetes mellitus), the adrenal	
CC	glands (e.g. hirsutism), ovaries, the thyroid (e.g. hyperthyroidism), the	
CC	hypothalamus and testes (e.g. vanishing testes syndrome), many examples	

CC of diseases and disorders are given in the specification. The
 CC present sequence represents an endocrine antigen of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 85 AA;

AAU20199 Length: 85 May 30, 2002 09:25 Type: P Check: 7019 ..

1 GCLVQNNQC SSSHLILRL ESQRPALARC LLEVFPSLE TLFIHWPE

51 QSEMAIPCCR GGNEMWLSRW MRFYRIKGS GGLRQ

!!AA_SEQUENCE 1.0

ID AAU27511 standard; Protein; 189 AA.

XX AC AAU27511:

XX DT 18-DEC-2001 (first entry)

XX DE Human G-Protein Coupled Receptor (GPCR) polypeptide #39.

XX KM Human: G-protein coupled receptor; GPCR; mental disorder; schizophrenia;
 KM attention deficit disorder; anxiety; depression; bipolar disorder;
 KM neurological disorder; Huntington's disease; dementia; obesity; anorexia;
 KM metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis;
 KM type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;
 KM cardiomypopathy; atherosclerosis; human immunodeficiency virus; HIV;
 KM viral infection; immunostimulant; neuroleptic; nootropic; tranquilizer;
 KM antidepressant; anorectic; gene therapy; antiviral; cardiant; alopecia;
 KM neuroprotective; cytostatic; antiparkinsonian; antidiabetic; psoriasis;
 KM hyperproliferative disorder; hormonal disorder; inflammatory disorder;
 KM Crohn's disease.

XX KM Homo sapiens.

XX OS WO200162798-A2.

XX PN 30-AUG-2001.

XX PD 23-FEB-2001; 2001WO-US05985.

XX PF 25-FEB-2000; 2000US-0184939.

XX PR 25-FEB-2000; 2000US-0184940.

XX PR 25-FEB-2000; 2000US-0184941.

XX PR 25-FEB-2000; 2000US-0184942.

XX PR 25-FEB-2000; 2000US-0185042.

XX PR 25-FEB-2000; 2000US-0185044.

XX PR 25-FEB-2000; 2000US-0185208.

XX PR 13-MAR-2000; 2000US-0184643.

XX PA (PHAA) PHARMACIA & UPJOHN CO.

XX PI Lind P, Parodi LA, Vogeli G, Wood LS;

XX DR WPI, 2001-582042/65.

XX DR N-PSDB; AAS43454.

XX PT Isolated nucleic acid molecules encoding G protein-coupled receptors

XX PT termed nGPCR-x, useful in the treatment and diagnosis of viral

XX PT infections, cancers and mental disorders (e.g. Parkinson's disease and

XX PT schizophrenia) -

XX PT Claim 31; Page 76; 21pp; English.

XX PS Sequences AAU27473-AAU27549 represent human G-protein coupled receptor

CC (GPCR) polypeptides of the invention. The proteins and their associated

CC DNA sequences can be used to identify compounds which bind to GPCR

CC polypeptides and in screening for compounds that modulate GPCR activity.

CC CC By screening a human subject for the presence of mutations in GPCR DNA, a

CC GPCR-related disorder or a genetic predisposition can be diagnosed. The

CC sequences can also be used for treatment and prevention of mental
 CC disorders such as schizophrenia, attention deficit disorder, anxiety,
 CC depression, dementia and bipolar disorder, neurological disorders such as
 CC Huntington's disease, Parkinson's disease and Tourette's syndrome,
 CC metabolic disorders such as obesity, anorexia and type 2 diabetes,
 CC cardiovascular disorders such as thrombosis, myocardial infarction,
 CC cardiomypopathy and atherosclerosis, hyperproliferative disorders such as
 CC psoriasis, hormonal disorders such as alopecia, inflammatory disorders
 CC such as Crohn's disease, viral infections caused by HIV and cancers.

XX Sequence 189 AA;

AAU27511 Length: 189 May 30, 2002 09:25 Type: P Check: 6166 ..

1 EVAMSRGSGE QTOLDNIMGL VFYIKVTRK MSQRPIIFCR VSIDYLLIQ

51 KAMKSLVEIE IQATHVRNCR APKYFRNVLS NKSIIKCSQP DIFGLNCRK

101 TDKKCPPLHK IITAKPVHHL TGECNKNVT ETRHGFANDW DTKLPGRIPS

151 RFVVTKKYQQ NSYRVLCFM SYWPFQFPFG GHDKGVTEL

!!AA_SEQUENCE 1.0

ID AAU29324 standard; Protein; 280 AA.

XX AC AAU29324:

XX DT 18-DEC-2001 (first entry)

XX DE Human PRO polypeptide sequence #301.

XX KM PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
 KM dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
 KM blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
 KM adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.

XX KM Homo sapiens.

XX OS WO200168848-A2.

XX PN 20-SEP-2001.

XX PD 28-FEB-2001; 2001WO-US06520.

XX PF 01-MAR-2000; 2000WO-US05601.

XX PR 02-MAR-2000; 2000WO-US05841.

XX PR 03-MAR-2000; 2000US-187202P.

XX PR 06-MAR-2000; 2000US-186968P.

XX PR 14-MAR-2000; 2000US-189328P.

XX PR 15-MAR-2000; 2000US-189328P.

XX PR 15-MAR-2000; 2000WO-US06884.

XX PR 21-MAR-2000; 2000US-190828P.

XX PR 21-MAR-2000; 2000US-191007P.

XX PR 21-MAR-2000; 2000US-191048P.

XX PR 28-MAR-2000; 2000US-192655P.

XX PR 29-MAR-2000; 2000US-193032P.

XX PR 29-MAR-2000; 2000US-193033P.

XX PR 30-MAR-2000; 2000WO-US08439.

XX PR 04-APR-2000; 2000US-194449P.

XX PR 04-APR-2000; 2000US-194647P.

XX PR 11-APR-2000; 2000US-195975P.

XX PR 11-APR-2000; 2000US-196000P.

XX PR 11-APR-2000; 2000US-196187P.

XX PR 11-APR-2000; 2000US-196690P.

XX PR 11-APR-2000; 2000US-196820P.

XX PR 18-APR-2000; 2000US-198121P.

XX PR 18-APR-2000; 2000US-198385P.

XX PR 25-APR-2000; 2000US-199397P.

XX PR 25-APR-2000; 2000US-199550P.

XX PR 25-APR-2000; 2000US-199654P.

XX PR 03-MAY-2000; 2000US-201515P.

XX PR 17-MAY-2000; 2000WO-US13705.

PR 22-MAY-2000; 2000MO-US14042.
 PR 30-MAY-2000; 2000MO-US14941.
 PR 02-JUN-2000; 2000MO-US15264.
 PR 05-JUN-2000; 2000US-209832P.
 PR 28-JUL-2000; 2000MO-US20710.
 PR 22-AUG-2000; 2000US-0644848.
 PR 24-AUG-2000; 2000MO-US23328.
 PR 08-NOV-2000; 2000MO-US30952.
 PR 01-DEC-2000; 2000MO-US32678.
 PR 20-DEC-2000; 2000MO-US34956.
 XX
 PA (GENTH) GENENTECH INC.
 XX
 PI Baker KP, Chen Y, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 XX
 DR WPI: 2001-602746/68.
 DR N-PSDB; AAS46225.
 XX
 PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 PT presence of tumours, such as prostate and breast tumours, in mammals and
 PT to screen for modulators of the compounds -
 XX
 PS Claim 11; Fig 602; 774pp; English.
 XX
 CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
 CC The PRO polypeptides and their associated nucleic acids can be used to
 CC detect the presence of a tumour in a mammal by comparing the level of
 CC expression of a PRO polypeptide in a test sample of cells from the animal
 CC and a control sample of normal cells, whereby a higher level of
 CC expression in the test sample indicates the presence of a tumour in the
 CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
 CC and rabbits but are preferably human. The polypeptides can be used to
 CC stimulate tumour necrosis factor (TNF) alpha release from human blood.
 CC CC when contacted with it. A specific polypeptide can be used to stimulate
 CC the proliferation or differentiation of chondrocyte cells. The PRO
 CC proteins can be used to determine the presence of tumours and also
 CC susceptibility to tumour development, particularly adrenal, lung, colon,
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders.
 XX
 SQ Sequence 280 AA;
 AAU29324 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..
 1 MQATYSTRD MDDDDGDTM SLHSQASRT RHPEPRTEH RAPSTMRPV
 51 ALTLITLCLV LLIGLALGL LFFOYYOLSN TGOOTISOME ERLGNTSOEL
 101 QSLAVQNIKL AGSLQHYAEK LCRELXNKAQ AHRCSPCTEQ WKWHDNCYQ
 151 FYKSKSKMED CKYFCLESENS TMLKINKQED LEFASQSYS EFFSYTGL
 201 LRPDSGKAWL WMDGTPFTSE LFHIIIDVTS PRSRDCAVAIL NGMIFSKDCK
 251 ELKRCVCERR AGWYKPESLH VPPELIGEGD
 !!AA_SEQUENCE 1.0
 ID AAU32368 standard; Protein: 273 AA.
 XX
 AC AAU32368;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #2859.
 XX
 KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 XX

PN WO200179449-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001MO-US08656.
 XX
 PR 18-APR-2000; 2000US-0552929.
 PR 26-JAN-2001; 2001US-0770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 PI WPI: 2001-611725/70.
 DR
 XX
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 XX
 PS Claim 20; Page 593-594; 765pp; English.
 XX
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemia. AAU29510-AAU3304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX
 SQ Sequence 273 AA;
 AAU32368 Length: 273 May 30, 2002 09:25 Type: P Check: 4809 ..
 1 VKGPGVKAE RFEXRMTAKH CALSLVGEPI MYPEINRPLK LLHOCKTSSP
 51 LVTNAQFPAA IRNLEPYQL YVRVASTKD SLKKIDRPLF KDFWQRFIDS
 101 VKALAVKQOR TVYRLTLVKA WNVESLQAYA QUGSLGNDF IEVGVYYCR
 151 ESSASSLTMA HVPWHEEVQ FVRELVDLIP EYEIACEHEH SNCLLIARHK
 201 FKIGEMWTM IDYRFOELI QEYEDSGSK TFSADYMAR TPWMAFGAS
 251 ERGFPKPTR HQRKNSKAI SGC
 !!AA_SEQUENCE 1.0
 ID AAM55526 standard; Protein: 27 AA.
 XX
 AC AAM55526;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 27631.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX

PF 30-JAN-2001; 2001MO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 27631; 650bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 27 AA;
AAM5556 Length: 27 May 30, 2002 09:25 Type: P Check: 8372 ..

1 SLKKIDRPLF KDFWQFLDS IKALAVK

!!AA_SEQUENCE 1.0

ID AAM67910 standard; Protein; 27 AA.

AC AAM67910;

DT 06-NOV-2001 (first entry)

XX Human bone marrow expressed probe encoded protein SEQ ID NO: 28216.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KM microarray; cancer; leukaemia; lymphoma; myeloma.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001MO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 28216; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 27 AA;
AAM67910 Length: 27 May 30, 2002 09:25 Type: P Check: 8372 ..

1 SLKKIDRPLF KDFWQFLDS IKALAVK

!!AA_SEQUENCE 1.0

ID AAM79021 standard; Protein; 377 AA.

AC AAM79021;

DT 06-NOV-2001 (first entry)

XX Human protein SEQ ID NO 1683.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

XX nervous system disorder; arthritis; inflammation.

OS Homo sapiens.

PN WO200157190-A2.

PD 09-AUG-2001.

PF 05-FEB-2001; 2001MO-US04098.

PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0663561.

PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Dymnac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wehrman T, Goodrich R;

DR WPI; 2001-476283/51.

DR N-PSDB; AAK52154.

XX Nucleic acids encoding polypeptides with cytokine-like activities,

PT useful in diagnosis and gene therapy -

XX Claim 20; Page 4030-4031; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78333-AAK80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX Sequence 377 AA;
SQ AAM79021 Length: 377 May 30, 2002 09:25 Type: P Check: 5679 ..
1 MOPPPPGGLG DCLRDWEDLQ ODFONIQETH RLYRLKLEEL TKLONNCTSS
51 ITRQKKRLQE LALALKKCR SLPAEAGAA QELENMKER QGLPDEMEAY
101 LPRKNGLYLS LVIGNNVTL LSKQAKFAK DEYEKFKLYL TITILISFT
151 CRFLNSRYT DAAENFLVW YCTLTIRFS ILINNSRIK GWWVFHHYVS
201 TFLSGVMTW PDGLMYOKFR NOFLSFSMTQ SFVQFLQYXY QSGCLYLRBA
251 LGERHTDUL VEGFQSWMMR GLTEFLPLFL FGHFWQLFNA LTLFNLAQDP
301 QCKEMQVIMC GEPFLLEFLG NEFTTLRVVH QLNHAAEBA PGAGPSPEEM
351 QTLPPSRGRG GRTGAGEPDE RARPRLL
11AA-SEQUENCE 1.0
ID AAM80005 standard; Protein: 406 AA.
XX
AC AAM80005;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3651.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN MO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
DR N-PSDB: AAK53138.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
XX useful in diagnosis and gene therapy -
XX
XX Claim 20; Page 407; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 406 AA;
XX
XX AAM80005 Length: 406 May 30, 2002 09:25 Type: P Check: 7978 ..
1 RSMAYVKRCK NNMCPNRGLH DGPEPCMLH AAGTSAVQA RGLQPSQRS
51 RPRVGLAFTA LAYGRAPTP LSRIGWAMP PRPGVLGDL RDWEDLQDP
101 ONIQVSAAD AGSPPSRVSL AOGQSGSGSPG CKPSLPAEAE GAAQELENM
151 KERQGLFFDM EAYLPKKNGL YLSLVGNVH VTLSKQAKF AYKDEYEXFK
201 LYLTILILILI SFTCRFLNS RVTDAFNFLL LVMYCTLTIL RESILINNS
251 RIKGWWVFHH VSTFLSGVM LTPWDGLMYO KFRNPLSFS MYQSFVDFLQ
301 YYYQSGCLYR LRALGERHTM DLTEGFGQSW MMRVLTFLPL FLFGHEFWQL
351 FNALTLFNIA QDPCKEMOV LMGFPFLLL FLGNFTTLIR VVHKHFSQR
401 HGSKKD
11AA-SEQUENCE 1.0
ID AAM84808 standard; Protein: 36 AA.
XX
AC AAM84808;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen SEQ ID NO:12401.
XX
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX
OS Homo sapiens.
XX
PN MO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234297.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0241060.
 PR 20-OCT-2000; 2000US-0241121.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246509.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-483426/52.
 DR N-PSDB; AAK57589.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS
 XX
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SO Sequence 36 AA;

AA04808 Length: 36 May 30, 2002 09:25 Type: P Check: 1969 ..

1 RPVIMCGYCG LIFCCCESSF WKLFRGHX VSXSL

11AA_SEQUENCE 1.0

ID AA04808 standard; Protein: 280 AA.

AC AA04808;

DT 06-NOV-2001 (first entry)

DE Human polypeptide, SEQ ID NO: 3297.

KW Human: full length cDNA; cDNA synthesis; oligo-capping.

OS Homo sapiens.

PN EP1130094-A2.

PD 05-SEP-2001.

PE 07-JUL-2000; 2000EP-0114089.

PR 08-JUL-1999; 99JP-0194486.

PR 11-JAN-2000; 2000JP-0118774.

PR 02-MAY-2000; 2000JP-0183765.

PA (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

DR WPI: 2001-524255/58.

DR N-PSDB; AA04808.

PT 830 Primers useful for synthesizing full length cDNA clones and their

PT use in genetic manipulation -

PS Claim 8; SEQ ID NO 3297; 1380bp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA

CC clones. 830 cDNA molecules encoding a human protein have been

CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA

CC molecules have been determined. Primers for synthesizing the full length

CC cDNA are useful for clarifying the function of the protein encoded by

CC the cDNA. The full length clones were obtained by construction of full

CC length enriched cDNA libraries that were synthesized by the oligo-capping

CC method. The primers enable the production of the full length cDNA easily

CC without any special methods. The present sequence is a polypeptide

CC encoded by a full length human cDNA of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 280 AA;

AA04808 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..

1 MOAKYSTRD MLDDGDTM SLHQAASAT RHPRRTEH RAPSSTWRPV

51 ALTLTLLV LTLGALGL LFFQYQLSN TGQDTISQME ERLGTSQEL

101 QSLQVQNK LAGSLQHAER LCRELYNKAG AHRCSPCTEQ WKMGDNQYQ

151 FYKDSKSWED CKYFCLSENS TWLKNQED LEFAASQSY EEFYSYWTGL

201 LRPDSGKAML WMDGTPFTE LFHIIIDVTS PRSRDCAVAL NGMIFSKDCK

251 ELKRCVCERR AGMWKPESLH VPPEITGEGD

11AA_SEQUENCE 1.0

ID AA001590 standard; Protein: 113 AA.

AA001590;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 15482.

KW Human: cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PE 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

PI WPI: 2001-514838/56.

DR N-PSDB; AA181521.

PT Isolated nucleic acids and polypeptides, useful for preventing

PT diagnosing and treating e.g. leukaemia, inflammation and immune

PS Claim 20; SEQ ID NO 15482; 1399bp + Sequence listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and

CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 113 AA;

AA001590 Length: 113 May 30, 2002 09:25 Type: P Check: 9232 ..

1 NNFLQLKFP PKKRTTDSY KAPRPFEPF PSQGVFPPT LFFGPPGFP

51 PPPPLNPPPG FFCFWAPLAK FFFPYPGXT WWSLQGPRLR FFFFFFXQGV

101 SLCHPGWSSS AQS

11AA_SEQUENCE 1.0

ID AA047415 standard; Protein: 585 AA.

AC AA047415;

DT 17-OCT-2001 (first entry)

DE Plant phosphate transporter, LePT3.

KW Plant; phosphate transporter; LePT3; solanaceae; tomato;

KW phyto-remediation; pollution; phosphate-based fertilizer;

OS Lycopersicon esculentum.

XX Key Location/Qualifiers

XX Domain 18..37

FT /note= "Membrane spanning domain I"

FT Domain 71..92

FT /note= "Membrane spanning domain II"

FT Domain 100..120

FT /note= "Membrane spanning domain III"

FT Domain 124..144

FT /note= "Membrane spanning domain IV"

FT Domain 171..191

FT /note= "Membrane spanning domain V"

FT Domain 211..232

FT /note= "Membrane spanning domain VI"

FT Modified-site 243..245

FT /note= "Caseln kinase II phosphorylation site"

FT Domain 292..312

FT /note= "Membrane spanning domain VII"

FT Domain 346..368

FT /note= "Membrane spanning domain VIII"

FT Domain 372..390

FT /note= "Membrane spanning domain IX"

FT Domain 399..419

FT /note= "Membrane spanning domain X"

FT Domain 441..462

FT /note= "Membrane spanning domain XI"

FT Domain 472..491

FT /note= "Membrane spanning domain XII"

FT Modified-site 498..501

FT /note= "Protein kinase C phosphorylation site"

XX

PN MO200155299-A2.

XX

XX 02-AUG-2001.

PD

XX

PF 26-JAN-2001; 2001MO-IL00075.

XX

XX 27-JUN-2000; 2000IL-0134258.

PR

XX

PA (YEDA) YEDA RES & DEV CO LTD.

PI

XX

PI Levy A;

XX

XX WPI; 2001-488780/53.

DR

XX

DR N-PSDB; AAH43180.

DR

XX

XX Phosphate transporter gene encoding a phosphate transporter, LepT3,

PT useful for generating plants which can be cultivated on

PT phosphate-deficient soil, and for phyto-remediation of phosphate

PT polluted soils -

PS

PS Claim 1; Fig 1; 59pp; English.

XX

XX This sequence shows a novel plant phosphate transporter, LepT3.

XX LepT3 is useful for increasing the uptake of phosphate by a plant

XX (belonging to the family solanaceae, preferably a tomato plant).

XX LepT3 cDNA is useful for producing a plant over expressing LepT3

XX which is useful for effecting phyto-remediation of an area polluted

XX with phosphate (a portion of which is radioactive) which involves:

XX (a) providing a plant (which belongs to the family solanaceae,

XX preferably a tomato plant) over expressing LepT3 to facilitate uptake

XX and concentration of phosphate within the plant cells;

XX (b) planting the plant in the area polluted with phosphate;

XX (c) following a time period, in which at least a fraction of the

XX phosphate in the area has been accumulated in the plant, harvesting the

XX plant, thereby removing at least the fraction of the phosphate from the

XX area; and

XX (d) optionally repeating steps the second and third steps until a

XX sufficient amount of the phosphate has been removed from the area.

XX Inactivating expression of the endogenous phosphate transporter gene

XX generates plants suitable for growth under high phosphate conditions.

XX LepT3 when expressed in plants increases the ability of the plants to

CC absorb phosphate in soil, either for purposes of phyto-remediation, or

CC as a means of reducing the need of phosphate-based fertilizers.

CC Therefore LepT3 DNA is useful for generating plants which can be

CC cultivated on phosphate-deficient soil and for phyto-remediation of

CC phosphate polluted soils. Plants which accumulate high concentrations

CC of phosphate due to LepT3 serve as an excellent phosphate sink in

CC phosphate polluted areas.

XX

XX Sequence 585 AA;

XX

AA047415 Length: 585 May 30, 2002 09:25 Type: P Check: 6481 ..

1 MASDNLVVLN ALDPTARTQWY HTVAIIAGM GFPTDAVDLF CITTISKLLG

51 RLYYDEPTTH APQKLEPHVAN NWVIGVALVG TLSQGLVFGW LQDKLGRKKV

101 YGTLILMVL CALCSGLSLG YSAKGVICTL CFFEFWLGFG IGGDYPLSAT

151 IMSEYANKAT RGAFTAAVFA MQGVGIFAG LVSHIISKLF LMYEGEPPN

201 VDEILSTEPQ ADYWRIVLM LGALPALLTY YWRMKMETG RYTAIIEGNA

251 KQAINNGKV LDIEIQESD KLAQFKANE YSLLSNEFFQ RHGLHLIGTM

301 STWFLLDIAP YSONLTQKDI FPVWGLTSNA NTISALEMP ETSRANFVIA

351 LFGTFPGYWF TWFFIEKIGR FRIQLMGFFM MSVFMAIIGV KYDYLSKKEH

401 KMTFATLYGL TFFANFGPN STFVLPAEL FPRVRSTCH ALSAASKGAG

451 AMISAFCIQO YQDGNVHKX QDSYTFVCY KYGILLPFL VYETKRSLSE

501 EITGEDGGQM RHRXRLANIS VSIKTTGWEY YKFONFLVIH HVALCESIFN

551 YLYEXELNHY EXPQYNKLVY TNSIVXSLHR HFYLF

IIAA_SEQUENCE 1 0

ID AAM15727 standard; Protein; 27 AA.

XX

AC AAM15727;

XX

DT 12-OCT-2001 (first entry)

XX

DE Peptide #2161 encoded by probe for measuring cervical gene expression.

XX

KM Probe: human; microarray; gene expression; cervical epithelial cell;

XX

OS Homo sapiens.

XX

PN MO200157278-A2.

XX

PD 09-AUG-2001.

XX

XX

PF 30-JAN-2001; 2001MO-US00670.

XX

XX 04-FEB-2000; 2000US-0180312.

PR

XX 26-MAY-2000; 2000US-0207456.

PR

XX 30-JUN-2000; 2000US-0608408.

PR

XX 03-AUG-2000; 2000US-0632366.

PR

XX 21-SEP-2000; 2000US-0234687.

PR

XX 27-SEP-2000; 2000US-0236359.

PR

XX 04-OCT-2000; 2000GB-0024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

XX WPI; 2001-488901/53.

DR

XX

PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human cervical epithelial cells -

XX

PS Claim 27: SEQ ID No 20553; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENP: see AAI10068-AA128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at http://wipo.int/publ/published_pct_sequences.
XX
XX Sequence 27 AA:

AA15727 Length: 27 May 30, 2002 09:25 Type: P Check: 8372 ..

1 SLKIDRPLF KDFWQOFLDS LKALAVK

IIAA_SEQUENCE 1.0
ID AAM28232 standard; protein; 27 AA.

AC AAM28232;

DT 17-OCT-2001 (first entry)

DE Peptide #2269 encoded by probe for measuring placental gene expression.

KW Probe: microarray; human; placenta; antenatal diagnosis;
KM genetic disorder.

OS Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -

PS Claim 27: SEQ ID No 28501; 654bp; English.

CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI1315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.

XX Sequence 27 AA:

AA28232 Length: 27 May 30, 2002 09:25 Type: P Check: 8372 ..
1 SLKIDRPLF KDFWQOFLDS LKALAVK
IIAA_SEQUENCE 1.0

ID AAM39779 standard; protein; 80 AA.
XX
XX AAM39779;
AC
XX 22-OCT-2001 (first entry)
DT
XX Human polypeptide SEQ ID NO 2924.

DE Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT.

PI WPI; 2001-442253/47.

PI N-PSDB; AAI58935.

PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -

PS Example 4; SEQ ID NO 2924; 10078bp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AA42213) with noctropic.

CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localized neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression.

CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX Sequence 80 AA:

AA39779 Length: 80 May 30, 2002 09:25 Type: P Check: 8327 ..

1 MFLCAQEWLIT LGINVPLIFY HFMRXFHCPA DSSELAYDPP VVNNADTLISY
51 CQKRAMCKLA FYLLSFYYL YCMITLVSS

IIAA_SEQUENCE 1.0
ID AAM39876 standard; protein; 531 AA.

```

XX AC AAM39876;
XX AC
XX DT 22-OCT-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 3021.
XX DE
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX KW peripheral nervous system; neuropathy; central nervous system; CNS;
XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX KW leukemia.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PE 26-DEC-2000; 2000MO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX DR WPI: 2001-442253/47.
XX DR N-PSDB; AA159032.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders
XX PT such as central nervous system injuries -
XX PS Example 4; SEQ ID NO 3021; 10078bp; English.
XX CC The invention relates to human nucleic acids (AA157798-AA161369) and
XX CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: Immune system suppression,
XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC assays for receptor activity, arthritis and inflammation, leukemias and
XX CC C.N.S disorders.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification.
XX CC
XX CC Sequence 531 AA:
XX CC
AAM39876 Length: 531 May 30, 2002 09:25 Type: P Check: 4914 ..
1 MADVLSTLRQ YNIOKKEIYV KGDEYIEGF SWPRKNVNTNY VMWGCKEBO
51 PREYTLDTST LELNNVHLIS HPYVYRRAT ENIPVARRPD RDDLGLYING
101 EASTSASIDR SAPLEIGLOR STOVKRADE VLAEAKRPRI EDECVRLDK
151 ERLAARLEGH KEGIVOTEOI RSLSEAMSVE KIAAIRAKIM AKKRSTIKTD

```

```

11AA_SEQUENCE 1.0
ID AAM40139 standard; Protein: 258 AA.
XX AC
XX AC AAM40139;
XX DT 22-OCT-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 3284.
XX DE
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX KW peripheral nervous system; neuropathy; central nervous system; CNS;
XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX KW leukemia.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PE 26-DEC-2000; 2000MO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX DR WPI: 2001-442253/47.
XX DR N-PSDB; AA159295.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders
XX PT such as central nervous system injuries -
XX PS Example 5; SEQ ID NO 3284; 10078bp; English.
XX CC The invention relates to human nucleic acids (AA157798-AA161369) and
XX CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: Immune system suppression,
XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

```

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 258 AA:

AA040139 Length: 258 May 30, 2002 09:25 Type: P Check: 2511 ..

1 MTVKKCALSL VGEPIIMPEI NREKLILHQC KISSPLVTNA QPFAIRNLE
 51 PYTQDYVSVD ASTRKDSLKRI DRPLFKDFWQ RFLDSLKALA VKQORTVYRL
 101 TLVKAMNVDL LQAYALVSL GNPDPFIEVKG VTYCGESSAS SLTMAHVPWH
 151 EEVQGFVHEL VDLIPEVEIA CEHEHSNCLL IAHKKKIGG EEWYTIIDYNR
 201 FQELIQEED SGGSKTFESAK DYMARPPHMA LFGASERGFQ PKDTRHQRKN
 251 KSKAISGC

!!AA_SEQUENCE 1.0
 ID AAM41565 standard; Protein: 104 AA.

AC AAM41565;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6496.

KW Human; nocotropic; immunosuppressant; cytosstatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PE 26-DEC-2000; 2000MO-US34263.

PF 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodlich R, Drmanac RT;

PI WPI: 2001-442253/47.

DR N-PSDB: AAI60721.

PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -

PS Example 2; SEQ ID NO 6496; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AA42213) with nocotropic,
 CC immunosuppressant and cytosstatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 104 AA:

AA041565 Length: 104 May 30, 2002 09:25 Type: P Check: 4766 ..

1 LRAVSLPLIQ LVLPYSIHS LFCIMFCAQ EMTTGLNVP LLEFHWRYF
 51 HCPADSEELA YDPPVYMNAD TLYSCQKEAW CKLAFYLLSF FYIYCMITY
 101 LVSS

!!AA_SEQUENCE 1.0
 ID AAM41925 standard; Protein: 207 AA.

AC AAM41925;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6856.

KW Human; nocotropic; immunosuppressant; cytosstatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PE 26-DEC-2000; 2000MO-US34263.

PF 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodlich R, Drmanac RT;

PI WPI: 2001-442253/47.

DR N-PSDB: AAI61081.

PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -

PS Example 2; SEQ ID NO 6856; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AA42213) with nocotropic,
 CC immunosuppressant and cytosstatic activity. The polynucleotides are useful

CC In gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 CC Sequence 207 AA;
 CC
 CC AAM41925 Length: 207 May 30, 2002 09:25 Type: P Check: 2688 ..
 CC
 CC 1 VTQLYVSIDA STKSLKKID RPLFKDFWQ FLDLSKALAV KQRTYRLT
 CC 51 LVKMNVDL QAYALVSLG NPDFLEKGV TYCGESSASS LTMHVPWHE
 CC 101 EVVQFVELY DLPEYEIAC EHEHNSCLLI AHRFKTIGE WMTWINYRF
 CC 151 QELIQEVEDS GGSKTSKAD VMARTPHML FGASERGFDP KDRRHQKKNK
 CC 201 SKAISGC
 CC
 CC !!AA_SEQUENCE 1.0
 CC ID AAU12400 standard; Protein: 280 AA.
 CC XX
 CC AC AAU12400;
 CC XX
 CC DT 24-OCT-2001 (first entry)
 CC XX
 CC DE Human PRO1131 polypeptide sequence.
 CC XX
 CC KM Human secretory and transmembrane; PRO: mammalian; cancer; lung;
 CC KM breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
 CC KM cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
 CC KM adipocyte; A-peptide; factor VIIA; gene therapy.
 CC XX
 CC OS Homo sapiens.
 CC XX
 CC PN MO200140466-A2.
 CC XX
 CC PD 07-JUN-2001.
 CC XX
 CC PF 01-DEC-2000; 2000WO-US32678.
 CC XX
 CC PR 01-DEC-1999; 99WO-US28301.
 CC PR 01-DEC-1999; 99WO-US28634.
 CC PR 02-DEC-1999; 99WO-US28551.
 CC PR 02-DEC-1999; 99WO-US28564.
 CC PR 02-DEC-1999; 99WO-US28565.
 CC PR 09-DEC-1999; 99US-0170262.
 CC PR 16-DEC-1999; 99WO-US30095.
 CC PR 20-DEC-1999; 99WO-US30911.
 CC PR 20-DEC-1999; 99WO-US30999.
 CC PR 30-DEC-1999; 99WO-US31243.
 CC PR 06-JAN-2000; 2000WO-US00277.
 CC PR 06-JAN-2000; 2000WO-US00376.
 CC PR 11-FEB-2000; 2000WO-US03565.
 CC PR 18-FEB-2000; 2000WO-US04341.
 CC PR 18-FEB-2000; 2000WO-US04342.
 CC PR 22-FEB-2000; 2000WO-US04414.
 CC PR 24-FEB-2000; 2000WO-US04914.
 CC PR 24-FEB-2000; 2000WO-US05004.
 CC PR 01-MAR-2000; 2000WO-US05601.
 CC PR 20-MAR-2000; 2000WO-US07377.
 CC PR 21-MAR-2000; 2000WO-US07532.
 CC PR 30-MAR-2000; 2000WO-US08439.
 CC PR 17-MAY-2000; 2000WO-US13705.

PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 10-NOV-2000; 2000WO-US30873.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,
 PI Gerlitsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S,
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX
 DR WPI: 2001-408281/43.
 DR N-PSDB: AAS214472.
 DR
 XX
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical
 XX
 PS Claim 12; Fig 458; 813pp; English.
 XX
 CC AAU12172-AAU12446 represent novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 CC XX
 CC SO Sequence 280 AA;
 CC
 CC AAU12400 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..
 CC
 CC 1 MQAKYSTRD MLDGDDPTM SLHSQASATP RHPEPRTEH RAPSSTWRPV
 CC 51 ALTLITLCLIV LILGLAALGL LFFQYYQLSN TGQDTISME ERLGNTSOEL
 CC 101 OSLOVONIKL AGSLQHYAEK LCRELKYNKAG AHRCSPTCEQ WKHGDNCYQ
 CC 151 FYKDSKSWED CKYFCLESENS TMLKINKORD LEFAASQSYS EFPYSWTGL
 CC 201 LRPSGKRWL WMDGTPTSE LFHIIIDVYS PRSNDCAVIL NGMIFSDCK
 CC 251 ELKRCVCERR AGMWKPSGLH VPPEITGEGD
 CC
 CC !!AA_SEQUENCE 1.0
 CC ID AAG92957 standard; Protein: 592 AA.
 CC XX
 CC AC AAG92957;
 CC XX
 CC DT 26-SEP-2001 (first entry)
 CC XX
 CC DE C glutamincum protein fragment SEQ ID NO: 6711.
 CC XX
 CC KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 CC KM organic acid synthesis.
 CC XX
 CC OS Corynebacterium glutamincum.
 CC XX
 CC PN EPI108790-A2.

XX 20-JUN-2001.
 PD 18-DEC-2000; 2000EP-0127688.
 PF 16-DEC-1999; 99JP-0377484.
 PR 07-APR-2000; 2000JP-0159162.
 PR 03-AUG-2000; 2000JP-0280988.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 PI WPI: 2001-376931/40.
 DR N-PSDB: AAH68176.
 DR
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT
 PS Claim 17; SEQ ID NO: 6711; 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 CC
 XX
 SO Sequence 592 AA;
 AAG92957 Length: 592 May 30, 2002 09:25 Type: P Check: 4387 ..
 1 MVSLEPLASL LTRRLATLKP ALKPRATHLAS LGAQYIAELV PGIRMSPNRR
 51 RILPANMGAG FIGAEIAMW ALSPSLPKP WWVTANLAV LQAVGHAAT
 101 GIHSLPRTN RRVSRKIYNA THATGATIL TTTVVGLRH RTQIRLIGOK
 151 NFGPRTIAG ISVGTGAYGA LITGELTQH SINEVKLLIE RLDPWISFI
 201 AAVSYITLTT LTLADRVILR RILHNSAIOA AHLNRMVPPG TEQWPEPERS
 251 GSPSYEYKMG AVSGQGRAYL SGGPRKDDII TYTRLSDETET HEPRITIGM
 301 VPGSLSDQV DLVIHEKRT GALRRDHIYI NNSTGTGMIT DWSAHTFEFL
 351 TGCNCVTISM QSYSLPSALS WYKNDGPIN AARMILDAVL HELDQLPTGS
 401 RPKLFLAGES LGAVGLAEVW GQVEKILGTA DCVLLSGAPR FSDAMNMLRT
 451 RRDASSSERL PVIDSGRHR FAGEPEHDM PATWOFPRMI VAQHASPDIV
 501 WMNAELFTRR PEWLKTPPKOD HODVPRRLW MPFVGMQOVA LDLETSTSV
 551 GGHCHNYHEE FIDYWALLD REVTPELRHS IAYWIRANHI KR
 IIA_SEQUENCE 1.0
 ID AAG92978 standard; Protein: 507 AA.
 XX AAG92978;
 AC
 XX 26-SEP-2001 (first entry)
 DT
 XX

DE C glutamicum protein fragment SEQ ID NO: 6732.
 XX
 KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.
 XX
 OS Corynebacterium glutamicum.
 OS
 PN EPI108790-A2.
 PN
 XX 20-JUN-2001.
 PD 18-DEC-2000; 2000EP-0127688.
 PF 16-DEC-1999; 99JP-0377484.
 PR 07-APR-2000; 2000JP-0159162.
 PR 03-AUG-2000; 2000JP-0280988.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 PI WPI: 2001-376931/40.
 DR N-PSDB: AAH68197.
 DR
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT
 PS Claim 17; SEQ ID NO: 6732; 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 CC
 XX
 SO Sequence 507 AA;
 AAG92978 Length: 507 May 30, 2002 09:25 Type: P Check: 2649 ..
 1 MGRMQSERLT PSDLVGCRVR QVQRINFPRI SPLPATMQR ARREVGLTEV
 51 LDRLEPQPK RGRIPFTRAD LNDDELAEF ETLLEIAAGD TLITGAVFTG
 101 TLEGVAMEVQ VDVLVRNPD SYMPVVSNN RVARBDPKT MGIATVTRLG
 151 LQPLELKAAT LRHHTIDGYR LTLALMGLE AGAARESSIG AVVGQDRDMA
 201 YLVDTTRYAP AARHALLTPA PTAPRRVKEC ATCRWPKEC PELKAADDIS
 251 LFLGDRADT YREKINTTT ALIDANLGEI SHVAAMPBAE IPVLRRAHT
 301 SAPRDEVIDE VDVEAYLDLG AYLMGAMDK TYIPFVMSD LGEAAEENF
 351 ARFWSMLKSR RDKAROGOT FGVFCTASNG EHNHMLSTAR RFECKVAGVP
 401 DEQETRSFIS SDQWMDFAV ARSOLVPGG LGLKQLAPAA GFHNEEDFA
 451 GEDSLHAYLI ASTAAPEPAE AARQQLSYN GDDCRATYAV RHMLRGART
 501 APVIGNT
 IIA_SEQUENCE 1.0

ID AAG74295 standard; Protein; 66 AA.
 XX AAG74295;
 AC
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen protein SEQ ID NO:5059.
 XX
 KW Human: colon cancer; colon cancer antigen; diagnosis; detection;
 XX colorectal carcinoma.
 OS Homo sapiens.
 XX
 PN MO200122920-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000MO-US26524.
 XX
 PR 29-SEP-1999; 99US-0157137.
 XX 03-NOV-1999; 99US-0163280.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX
 DR WPI; 2001-235357/24.
 DR N-PSDB; AAH33726.
 XX
 PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 XX useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 PS Claim 11: Page 6783-6784; 9803pp; English.
 XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated PS,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAH77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SO Sequence 66 AA;
 AAG74295 Length: 66 May 30, 2002 09:25 Type: P Check: 942 ..
 1 TKLHFGQGL GNXLIVKSCN TSVQVNIIGP CFPSCMHEL FPMHMGASQ
 51 WXLNLPVGIIG XTWACL
 IIAA_SEQUENCE 1.0
 ID AAG8196 standard; Protein; 160 AA.
 XX
 AC AAG8196;
 XX
 DT 11-SEP-2001 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 316.
 XX
 KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
 KW GENSET.

XX
 OS Homo sapiens.
 XX
 PN MO200142451-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 07-DEC-2000; 2000MO-IB01938.
 XX
 PR 08-DEC-1999; 99US-0169629.
 XX 06-MAR-2000; 2000US-0187470.
 XX
 PA (GENSET) GENSET.
 XX
 PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
 XX
 DR WPI; 2001-367870/38.
 DR N-PSDB; AAH64799.
 XX
 PT Full length GENSET human nucleic acids encoding potentially secreted
 XX proteins, useful in gene therapy and vaccination against a variety of
 XX diseases, and for diagnosis of those diseases -
 XX
 PS Claim 21: Page 841; 921pp; English.
 XX
 CC The invention relates to full length GENSET human nucleic acids encoding
 CC potentially secreted proteins. The nucleic acids and the polypeptides
 CC they encode may be used in the prevention, treatment and diagnosis of
 CC diseases associated with inappropriate GENSET gene expression. For
 CC example, they be used to treat disorders associated with decreased
 CC GENSET gene expression by rectifying mutations or deletions in a
 CC patient's genome that affect the activity of GENSET or by supplementing
 CC the patient's own production of GENSET polypeptides. Conversely,
 CC antisense nucleic acid molecules may be administered to down regulate
 CC GENSET expression by binding with the cells' own genes and preventing
 CC their expression. The sense and antisense nucleic acids may also be
 CC used as DNA probes in diagnostic assays to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples, and hence to
 CC determine which patients may be in need of restorative therapy.
 CC The GENSET polypeptides may be used as antigens in the production of
 CC antibodies and in assays to identify modulators (agonists and
 CC antagonists) of GENSET polypeptide expression and activity. The
 CC present sequence is a GENSET polypeptide of the invention.
 XX
 SO Sequence 160 AA;
 AAG8196 Length: 160 May 30, 2002 09:25 Type: P Check: 4447 ..
 1 MATFFAFCY MSLVLCAL IFFATWHITA FDELRTDFKS PIDQCPVHA
 51 RERLRNIERI CFLLRRLVLP EYSIHSLFCI MFLCAQEWLT LGINVPLFVY
 101 HFWRYFHCPA DSSELAYDPP VVMNPDTLSY CQKEAWCKLA FYLLSFFYTL
 151 YCMITVTVSS
 IIAA_SEQUENCE 1.0
 ID AAU03216 standard; Protein; 483 AA.
 XX
 AC AAU03216;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Fruit fly G protein coupled receptor, DmGPCR10.
 XX
 KW Fruit fly; G protein coupled receptor; DmGPCR10;
 KW human immunodeficiency virus; HIV; cancer; Parkinson's disease;
 KW diabetes; obesity; atherosclerosis; thrombosis; stroke; renal failure;
 KW inflammation; rheumatoid arthritis; autoimmune disorder;
 KW neurological disorder; schizophrenia; manic depression; dementia;
 KW severe mental retardation; dyskinesia; Huntington's disease;
 KW Tourette's syndrome.
 XX

OS Drosophila melanogaster.
 XX
 PN MO200131005-A2.
 XX
 PD 03-MAY-2001.
 XX
 PF 20-OCT-2000; 2000WO-US29002.
 XX
 PR 22-OCT-1999; 99US-0425676.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 PI Lowery DE, Smith VG, Kudlak TA, Larsen MJ;
 XX
 DR WPI; 2001-316333/33.
 DR N-PSDB; AAS05895.
 PT New Drosophila melanogaster GPCR nucleic acids and polypeptide useful
 PT for inducing an immune response, for identifying homologs and for
 PT treating e.g. diabetes, obesity and manic depression -
 XX
 PS Claim 29; Page 66; 110pp; English.
 XX
 CC The sequence is a fruit fly G protein coupled receptor, DmGPCR10.
 CC The proteins are useful for inducing an immune response against itself in
 CC a mammal. The nucleic acids are useful for identifying an animal homolog
 CC of DmGPCR, by screening databases or libraries. The compounds identified
 CC as binding partners or modulators of GPCR binding are useful for treating
 CC diseases in animals, and for control insects that are harmful or cause
 CC injury to plants or animals. Diseases treated include infections (e.g.
 CC viral and human immunodeficiency virus, HIV), cancer, pain, Parkinson's
 CC disease, hypotension, hypertension, diabetes, obesity, atherosclerosis,
 CC thrombosis, stroke, renal failure, inflammation, rheumatoid arthritis,
 CC autoimmune disorders, and psychotic and neurological disorders (anxiety,
 CC schizophrenia, manic depression, delirium, dementia, severe mental
 CC retardation, dyskinesias, Huntington's disease or Tourette's syndrome).
 CC The nucleic acids can be used for genetic mapping, and producing
 CC the GPCR. Anti-GPCR antibodies can be used in therapy, diagnostic assays
 CC and for modulating GPCR activity.
 XX
 XX Sequence 483 AA;
 SQ
 AAU03216 Length: 483 May 30, 2002 09:25 Type: P Check: 0 ..
 1 MYASLMDVQG TLAARLADSD GNGANDSGL ATGGLGEQ EGLALDGNH
 51 ASADGGIVPV VPVLDRETY IYVLYTLTF IVGVGNCTL VIIFRRISM
 101 RNIDNTYILS LALADLLVIL VCVPAVTIVY TQESWPEERN MCRISSEFKD
 151 ISIGSVFTL TALSGERYCA IYNPLRKIQT KPLVFTAVM IWLAILLGM
 201 PSVLFSDIKS YPVFTATGNM TIEVCSFRD PEYAKFMAG KALVYVLLPL
 251 SIIGALYIMM AKRLHMSARN MPEQOSMOS RTQARARLHV ARMYAVFVV
 301 FFICFPFPHV FELMYHEPYT AEDEDEFWN VLRIPLKLV RQRLGYCSG
 351 VFRQHPNRYL CCICVAKRPH LRQHSATGCM MDNTSVMSMR RSTYVGSTAG
 401 NLRASLHRNS NHGVGAGAGG VGGVGSGRY GSFHRDSMP LQHGNAHGG
 451 AGGSSSGIGA GGRTAAVSEK SFINRESCV MRY
 11AA SEQUENCE 1 0
 ID AAB93312 standard; Protein: 732 AA.
 XX
 AC AAB93312;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:12391.

KW Human; primer: detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 12391; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 XX Sequence 732 AA;
 SQ
 AAB93312 Length: 732 May 30, 2002 09:25 Type: P Check: 4753 ..
 1 MDPASDTWDL FSPPLISLMIN RFIYLGFAV SISLWICQI VIKQGNLQ
 51 EKSVPKAOD LMTNGVYSIQ EKDIFVGVK IFYSGOTGTA KGAFTVLAEA
 101 VTSIDLPAVI INLKEYDPD HLIEVTSKN VCVPLVATYT DGLPTESAEM
 151 FCKWLEASI DFRGKTYLK GMRYAIFGIG NSAYASHNK VGRNVDMWM
 201 MGAHRVMSR GEGDCDVVKS KHGSIADPR AMKTFISQL QALOKGERKK
 251 SCGCHCKKKG CESHQSGSEE REEGSHODE LHHRDTEEE PFESSSEEF
 301 GGEHQSLNS IYVEDIGKI MDHVKKERK KEQGEKSGL FRNNGRNEG
 351 ERRAMITPAL REALTKQGY LIGSHGVKL CRWTKMLNG RGGCYKHTFY

401 GIESHRCMET TPLSLACANKC VECMRHHTNP VGTEWRKMD QPEMILKEAI
 451 ENHONMTKQF KGVPGVKAER FEEGMTVKHC ALSLVGEPIM YPEINRPLKL
 501 LHOCKISSFL VTNAQFPALR RNLEPVTOLY VSVDASTRDS LKKIDRPLFK
 551 DEMQRFDSL KALAVKQORT YRRLTVKRW NVDELQAYAO VLSLGNDFI
 601 EVKGVTCGE SSASSLTMAH VPMHEEVQF VRELVDLPE YEACEHEHS
 651 NCLLIARHKE KIGGEWMTWI NYNRFQELIQ EYEDSGSKT FSAKDYMART
 701 PHMALFGASE RGFDPKDTIH QKKNKSKAIS GC

11AA_SEQUENCE 1.0
 ID AAB94345 standard; Protein; 561 AA.
 AC AAB94345;
 XX
 DT 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:14855.
 XX
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.
 XX
 PI EP1074617-A2.
 PN 07-FEB-2001.
 PD
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS
 XX
 PS Claim 8; SEQ ID 14855; 2537pp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 XX Sequence 561 AA;
 SO

AAB94345 Length: 561 May 30, 2002 09:25 Type: P Check: 6460 ..

1 MRYAVFGLGN SAYASHENKV GRNVKWLMM LGARVMSRG EGDQDVYKSK
 51 HGSLEADPRA WTKRFTSOLQ ALQKGERKKS CGGHCKKQKC ESHQHSSEER
 101 EEGSHEDDEL HHRDTEEREP FESSSEEEFG GEDHOSLNSI VDVEDLCKIM
 151 DTVKKERERK EQQEEKSGLF RNMGRNEDGE RRAMITPALR EALTQOQYQL
 201 IGSHSQVKLC MWTKSMRGR GGQYKRTFVG IESHRCMETT PSLAYANKCV
 251 FCMRHHTNPV GTEWRKMDQ PEMILKEALE NHONNIKQFK GVPGVKAERF
 301 EEGMTVKHCA LSLVGEPIMY PEINRFLKL HQCKISSFLV TNAQFPAEIR
 351 NLEPVTOLYV SVDASTRKDSL KRIDRPLFKD FWQRFDSLK ALAVKQORTV
 401 YRRLTVKRWV VDELQAYAO VLSLGNDFIE VKGVTCGES SASSLTMAHV
 451 PMHEEVQFV RELVDLPEY EIACEHEHSN CLLIARHKE IGGEWMTWIN
 501 YNRFQELIQE YEEDSGSKTF SAKDYMARTP HMALFGASER GFDPKDTIRHQ
 551 RKNKSKAISG C

11AA_SEQUENCE 1.0
 ID AAB60464 standard; Protein; 160 AA.
 XX
 AC AAB60464;
 XX
 DT 24-APR-2001 (first entry)
 DE
 XX
 XX Human cell cycle and proliferation protein CCYPR-12, SEQ ID NO:12.
 DE
 XX
 KM Cell cycle and proliferation protein; CCYPR; human; agonist;
 KM antagonist; gene therapy; detection; gene therapy;
 KM developmental animal disease model; immune disorder;
 KM developmental disorder; cell signalling disorder;
 KM cell proliferative disorder; cancer; tumour; anaemia;
 KM arteriosclerosis; asthma; allergy; diabetes mellitus;
 KM menstrual cycle disorder; bacterial infection.
 XX
 OS Homo sapiens.
 XX
 PN WO200107471-A2.
 PN
 PD 01-FEB-2001.
 PD
 XX
 PF 21-JUL-2000; 2000MO-US19948.
 XX
 PR 21-JUL-1999; 99US-0145075.
 PR 08-SEP-1999; 99US-0153129.
 PR 10-NOV-1999; 99US-0164647.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Hillman JL, Lal P, Tang YF, Yue H, Au-Young J, Bandman O;
 PI Azimzal Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
 XX
 DR WPI; 2001-112727/12.
 DR N-PSDB; AAF59601.
 XX
 PT Human cell cycle and proliferation proteins and polynucleotides are
 PT used to treat, diagnose and prevent immune, developmental and cell
 PT signalling disorders and cell proliferative disorders including cancer -
 XX
 PS Claim 1; Page 123; 205pp; English.

XX Sequences AAB60453-AAB60506 represent 54 human cell cycle and
 CC proliferation proteins (CCYPR), which are encoded by AAF59590-AAF59643.
 CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
 CC associated with decreased expression of functional CCYPR, while CCYPR
 CC antagonists are used to treat diseases or conditions associated with
 CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
 CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
 CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
 CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
 CC that specifically bind to CCYPR, and in drug screening methods to
 CC identify compounds that modulate the activity of CCYPR. CCYPR
 CC nucleotides can be used to generate transgenic animal models of human
 CC disease, and can be used in gene therapy in target cells with genetic
 CC abnormalities with respect to the expression of CCYPR for the
 CC treatment or prevention of a disorder associated with CCYPR.
 CC Diseases which can be diagnosed, treated and prevented using CCYPR
 CC proteins, nucleic acids, agonists or antagonists include immune,
 CC developmental and cell signalling disorders, and cell proliferative
 CC disorders including cancer. Specific examples of these disorders
 CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
 CC diabetes mellitus, disorders of the menstrual cycle and infections
 CC caused by bacteria.

XX Sequence 160 AA;

AAB60464 Length: 160 May 30, 2002 09:25 Type: P Check: 4282 ..

1 MAFTPAACY MSLVLCNAL IFFAIWHITA FDELRTDFRS PIDOCNPPIHA
 51 RERLRNIERI CELRLKLVLP EYSIHSLFCT MFLCAQEWLT LGLNVPLLFY
 101 HFWRFHCPA DSESLAYDPP VVMNADTLXY CQKAWCKLA FYLSFFYYL
 151 YCMITLVVSS

!!AA_SEQUENCE 1.0
 ID AAB60465 standard; Protein: 531 AA.

XX AAB60465;

XX 24-APR-2001 (first entry)

XX Human cell cycle and proliferation protein CCYPR-13, SEQ ID NO:13.

XX Cell cycle and proliferation protein: CCYPR; human; agonist;
 KW antagonist; gene therapy; detection; gene therapy;
 KW transgenic animal disease model; immune disorder;
 KW developmental disorder; cell signalling disorder;
 KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
 KW arteriosclerosis; asthma; allergy; diabetes mellitus;
 KW menstrual cycle disorder; bacterial infection.

XX Homo sapiens.

XX WO200107471-A2.

XX 01-FEB-2001.

XX 21-JUL-2000; 2000WO-US19948.

XX 21-JUL-1999; 99US-0145075.
 PR 08-SEP-1999; 99US-0153129.
 PR 10-NOV-1999; 99US-0164647.

XX (INCY-) INCYTE GENOMICS INC.

XX Hillman JL, Lai P, Tang YT, Yue H, Au-Young J, Bandhan O,
 PI Azimzal Y, Yang J, Lu DM, Baughn MR, Patterson C, Shah P;
 XX WPI: 2001-112727/12.

DR N-PSDB; AAF59602.

XX

PT Human cell cycle and proliferation proteins and polynucleotides are
 PT used to treat, diagnose and prevent immune, developmental and cell
 PT signaling disorders and cell proliferative disorders including cancer -
 PS Claim 1; Page 123-124; 205pp; English.

XX Sequences AAB60453-AAB60506 represent 54 human cell cycle and
 CC proliferation proteins (CCYPR), which are encoded by AAF59590-AAF59643.
 CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
 CC associated with decreased expression of functional CCYPR, while CCYPR
 CC antagonists are used to treat diseases or conditions associated with
 CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
 CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
 CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
 CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
 CC that specifically bind to CCYPR, and in drug screening methods to
 CC identify compounds that modulate the activity of CCYPR. CCYPR
 CC nucleotides can be used to generate transgenic animal models of human
 CC disease, and can be used in gene therapy in target cells with genetic
 CC abnormalities with respect to the expression of CCYPR for the
 CC treatment or prevention of a disorder associated with CCYPR.
 CC Diseases which can be diagnosed, treated and prevented using CCYPR
 CC proteins, nucleic acids, agonists or antagonists include immune,
 CC developmental and cell signalling disorders, and cell proliferative
 CC disorders including cancer. Specific examples of these disorders
 CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
 CC diabetes mellitus, disorders of the menstrual cycle and infections
 CC caused by bacteria.

XX Sequence 531 AA;

AAB60465 Length: 531 May 30, 2002 09:25 Type: P Check: 4914 ..

1 MADVLSVLQK YNIOKKEIYV KQDEVIFGEF SMPKNVKTNY VVMGCKRQSG
 51 PREYTLDSI LFLNNVHLS HPVYVRRPAT ENIPVVRPPD RKDLGLYING
 101 EASTSASIDR SAPLEIGLQK STQVKRADE VLAFAKKPRI EDEECVRLDK
 151 ERLARLEGH KEGIVQTEQI RSLSEAMSV KTAIAKAKIM AKKSTITTD
 201 LDDITAIKQ RSEVDAEYDV TRDIVSRERY WRTPTTILOS TGNKESKNIF
 251 AIIQSVKARE EGRAPEORPA PNAAPVPTL PTKQPIPAAY NRYQDERFKG
 301 KEETEGFKID TMGYTHGHTL KSVTEGASAR KQTPAAGVY PRPSQARPP
 351 PNOKKGSRTF IITIPATTS LITMLNAKDL LODLAFVPSD EKKQCGORE
 401 NETLIQRRKD QMOPGTAIS VTPVRYVVDQ PLKLPQDWD RVAVVFVQGP
 451 AMQFGWPMWL LPDGSVDIF AKIKAFHLKY DEVRLDPNVQ KMDVTYVLELS
 501 YAKRHLDPRV FLRFWETLDR YMWKHKSHLR F

!!AA_SEQUENCE 1.0
 ID AAB65251 standard; Protein: 280 AA.

XX AAB65251;

XX 02-APR-2001 (first entry)

XX Human PRO1131 (UNQ569) protein sequence SEQ ID NO:319.

KW Human; secreted and transmembrane protein; PRO; cytosolic;
 KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
 KW diagnostic assay.

XX Homo sapiens.

XX WO200073454-A1.

XX 07-DEC-2000.

PD

XX 30-MAR-2000; 2000MO-US08439.
PF 02-JUN-1999; 99MO-US12252.
XX 23-JUN-1999; 99US-0141037.
PR 07-JUL-1999; 99US-0143048.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 17-AUG-1999; 99US-0149396.
PR 15-SEP-1999; 99MO-US21090.
PR 08-OCT-1999; 99US-0158663.
PR 30-NOV-1999; 99MO-US28313.
PR 01-DEC-1999; 99MO-US28301.
PR 16-DEC-1999; 99MO-US30095.
PR 20-DEC-1999; 99MO-US30911.
PR 05-JAN-2000; 2000MO-US00219.
PR 06-JAN-2000; 2000MO-US00376.
PR 11-FEB-2000; 2000MO-US03565.
PR 18-FEB-2000; 2000MO-US04341.
PR 22-FEB-2000; 2000MO-US04414.
PR 24-FEB-2000; 2000MO-US04914.
PR 24-FEB-2000; 2000MO-US05004.
PR 02-MAR-2000; 2000MO-US05841.
PR 15-MAR-2000; 2000MO-US05884.
PR 20-MAR-2000; 2000MO-US07377.
XX
XX (GETH) GENENTECH INC.
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL,
PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ,
PI Grimaldi CJ, Gurney AL, Kijavlin IJ, Napier MA, Pan J, Paoni NF,
PI Roy MA, Stewart TA, Tamas D, Watanabe CK, Williams PM, Wood WI,
PI Zhang Z;
XX MPI; 2001-032160/04.
DR N-PSDB; AAF44218.
XX
PT PRO polynucleotides used to produce polypeptides used to target
XX bioactive molecules such as toxins, radiolabels or antibodies, to
XX specific cells, to cause targeted cell death -
XX
XX Claim 12; Fig 230; 935pp; English.
XX
PS The present invention describes human secreted and transmembrane PRO
CC proteins. The PRO proteins have cytostatic activity. The PRO proteins
CC can be used for targeted delivery of bioactive molecules, such as
CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
CC sequences, and their fragments, can be used as hybridisation probes, in
CC chromosomal and gene mapping, and in the generation of anti-sense RNA
CC and DNA. They may also be used to produce transgenic animals which are
CC used to develop and screen therapeutically useful reagents. The PRO
CC nucleotide and protein sequence can be used for tissue typing and in
CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.
XX
XX Sequence 280 AA;
SO
AAB65251 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..
1 MQAKYSSTRD MUDDGDDTTM SLHSQASATY RHPEPRRTEH RAPSSTWREY
51 ALRLTLTCLV LLIGLAALGL LFPOYQUSN TGDDTISQME ERLGNTSQEL
101 QSLQVQNIKL ASSLQHVARE LCRELYNKAG AHRCSPECTQ WKWGNQCYQ
151 FYDKSKWED CKYFCLSENS TMLKINKOED LEFAASQSYS EFFYSYWTGL
201 LRPDGKAWL WMDGTPTSE LFHIIIDVTS PRSHDCAAIL NCMIFSKDCK

251 ELKRCVCERR AGMYKPESEL VPPETLGEGD
IIAA_SEQUENCE 1.0
ID AAB50959 standard; Protein: 280 AA.
XX
AC AAB50959;
XX
DT 21-MAR-2001 (first entry)
XX
DE Human PRO1131 protein.
XX
KW Human; PRO; cytostatic; nootropic; neuroprotective; respiratory general;
KW antiinflammatory; antiangiogenic; immunosuppressive; immunostimulant;
KW PRO agonist; cancer; inflammatory disorder; immunological disorder.
XX
OS Homo sapiens.
XX
PN WO200073348-A2.
XX
PD 07-DEC-2000.
XX
PF 30-MAY-2000; 2000MO-US14941.
XX
XX 02-JUN-1999; 99MO-US12252.
PR 22-JUN-1999; 99US-0140650.
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 01-SEP-1999; 99MO-US20111.
PR 08-SEP-1999; 99MO-US20594.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99MO-US28313.
PR 01-DEC-1999; 99MO-US28634.
PR 02-DEC-1999; 99MO-US28551.
PR 16-DEC-1999; 99MO-US30095.
PR 20-DEC-1999; 99MO-US30999.
PR 06-JAN-2000; 2000MO-US00376.
PR 11-FEB-2000; 2000MO-US03565.
PR 18-FEB-2000; 2000MO-US04341.
PR 18-FEB-2000; 2000MO-US04342.
PR 02-MAR-2000; 2000MO-US05841.
PR 03-MAR-2000; 2000US-0187202.
PR 10-MAR-2000; 2000MO-US06319.
PR 15-MAR-2000; 2000MO-US06884.
PR 30-MAR-2000; 2000MO-US08439.
PR 17-MAY-2000; 2000MO-US13705.
XX
XX (GETH) GENENTECH INC.
PI Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W, Kabakoff RC,
PI Shelton DL, Smith V, Watanabe CK, Wood WI;
XX MPI; 2001-016509/02.
DR N-PSDB; AAC91561.
XX
PT Twenty eight nucleic acids encoding PRO polypeptides which are useful
XX for treating various tumors, e.g. breast cancer, and other
XX inflammatory, angiogenic and immunological disorders -
XX
XX Claim 31; Fig 18; 188pp; English.
XX
PS The present sequence is one of twenty eight novel PRO polypeptides. The
CC PRO polypeptides and their agonists, including antibodies, peptides, and
CC small molecule agonists, may be used to treat various tumors, e.g.,
CC cancers such as breast cancer, ovarian cancer, renal cancer, colorectal
CC cancer, uterine cancer, prostate cancer, lung cancer, bladder cancer,
CC central nervous system cancer, melanoma or leukemia. They are also
CC useful for treating other disorders such as neuronal, glial, astrocytal,
CC hypochalamic and other glandular, macrophagal, epithelial, stromal and
CC blastocoealic disorders, and inflammatory, angiogenic and immunological
XX disorders.
XX
XX Sequence 280 AA;
SO

AAB50959 Length: 280 May 30, 2002 09:25 Type: P Check: 2849 ..

1 MQAKYSTRD MLDDGDITM SIHSQASATP RHPERREH RAPSSTRPV

51 ALTLTLCLV LILGLALGI LFEOYYOLSN TGODTISOME ERLGNTSOEL

101 OSLOVONIKL AGSLQHAEK LCRELYNKAG AHRCSPCTEQ WKMHGNCYQ

151 FYKSKSMED CKYFCLENS TWLKINKQED LEFAASQSY S EFFSYWTGL

201 LRPDSGAWL MMDGPTPSE LFPHIIDVTS PRSRDCAVIL NGMIFSKCK

251 ELKRCVCERR AGWKPESELH VPPELLEGCD

11AA_SEQUENCE 1.0

ID AAU19776 standard; Protein: 394 AA.

XX AAU19776;

AC

XX

DT 06-DEC-2001 (first entry)

XX

DE Human novel extracellular matrix protein, Seq ID No 426.

XX

KM Human; secreted extracellular matrix protein; immunomodulatory;

KM Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiac; vascular;

KM cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cyostatic;

KM antialzheimers; immune/autoimmune disease; HIV infection; anaemia;

KM human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;

KM cancers; hyperproliferative disorder; breast neoplasia; melanoma;

KM Sezary syndrome; Gaucher's disease; neurological diseases;

KM Alzheimer's disease; Parkinson's disease; cardiovascular disorder;

KM cardiac arrest; tachycardia; angina; infection; corneal infections;

KM wound healing; immunogen; gene therapy; antisense; food additive.

XX Homo sapiens.

OS

XX WO20015368-A1.

PN

XX

PD 02-AUG-2001.

XX

PE 17-JAN-2001; 2001WO-US01348.

XX

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217486.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226688.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 DR N-PSDB; AAS31347.
 XX
 PT Nucleic acid molecules encoding human secreted extracellular matrix
 PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
 PT Alzheimer's and Parkinson's diseases and cancers -
 XX
 PS Claim 11: SEQ ID No 426; 577pp; English.
 XX
 CC The invention relates to isolated nucleic acid molecules encoding
 CC novel human secreted extracellular matrix proteins (SPs). The
 CC polynucleotides and proteins are used to prevent, treat a medical
 CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
 CC chickens or sheep. For example, disorders associated with decreased
 CC expression of SPs. The SP polynucleotide or a vector expressing them may
 CC be administered to treat diseases by gene therapy. Antisense molecules
 CC may be administered to down regulate expression of SPs by binding with
 CC the cells own genes and preventing their expression. The polynucleotides
 CC may also be used as DNA probes in diagnostic assays. The SPs may also be
 CC used as antigens to produce antibodies and to identify modulators
 CC (agonists and antagonists) of the SPs. The anti-(SP) antibodies and
 CC antagonists may also be used to down regulate expression and activity of
 CC SP and as diagnostic agents for detecting the presence of SPs in samples.
 CC The disorders include for example: immune/autoimmune diseases (e.g. HIV
 CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
 CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
 CC melanomas, neoplasms of the breast or liver, Secary syndrome and
 CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
 CC Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac
 CC arrest, tachycardia and angina), infections caused by bacteria, viruses
 CC and fungi and ocular disorders (e.g. corneal infections). Other uses
 CC include wound healing, maintenance of organs before transplantation,
 CC support of cell culture of primary tissues, modulation of for example
 CC differentiation of embryonic stem cells, mammalian characteristics
 CC (e.g. height and weight), catabolism, anabolism, energy storage,

CC mental state, biorhythms, cardiac rhythms, reproductive potential,
 CC hormonal levels appetite, memory and stress. The SPs can also be used as
 CC an food additive. Full details of uses diseases are disclosed in the
 CC specification. The present sequence represents a novel human secreted
 CC extracellular protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 394 AA:
 AAU19776 Length: 394 May 30, 2002 09:25 Type: P Check: 52 ..
 1 DRRPASGCP GLLRREPLLA TGAXAMQPP PGPGLDCIRD WEDLQDDFON
 51 IQETHRLYRL KLEELTKLQN NCTSSITRQK KRLQELALAL KKCPSPFAP
 101 AECAQAELEN QMKEQGLFE DMEAYLPKKN GLYISIVLGN VNVTLISKQA
 151 KFAVKDEYEK FKLYITITLI LISFTCRFLI NSRVTDAFN FLVWRYCTL
 201 TIRESLINN GSRKGMWVF HHVSTFLSG VMLTWPGLM YQFRNQFLS
 251 FSWQSFVQF LQYYQSGCL YRLRALGERH TMDLTVEGFQ SWMMRGITFL
 301 LPFLFFGHFW QLENALTFEN LAQDPQCKEW QGCAPOVSQS AAREQEGRL
 351 GLPLPAQGRF CPVCCGRGWE APLECACIRG SLFSLGYFG RCGP
 IIAA SEQUENCE 1.0
 ID ABB48814 standard; Protein; 510 AA.
 XX
 AC ABB48814;
 XX
 DT 05-FEB-2002 (first entry)
 XX
 DE Listeria monocytogenes protein #1518.
 XX
 KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.
 XX
 OS Listeria monocytogenes.
 XX
 PN W0200177335-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 11-APR-2001; 2001WO-FR01118.
 XX
 PR 11-APR-2000; 2000FR-0004629.
 XX
 PA (INSP) INST PASTEUR.
 XX
 PI Buchteser C, Frangeul L, Couve E, Rusniok C, Pailh H, Dehoux P;
 PI Dussurget O, Chehouani F, Nedjarl H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 XX
 DR WPI; 2002-010914/01.
 XX
 CC Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 CC and prevention of Listeria and related bacterial infections, and
 CC related polypeptides -
 XX
 PS Claim 6: SEQ ID No 1519; 192pp; French.
 XX
 CC The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of

CC It are useful for selecting probes and primers for detecting genes in L.
 CC monocytoenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytoenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytoenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by L.
 CC monocytoenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 510 AA:

ABBA8814 Length: 510 May 30, 2002 09:25 Type: P Check: 2157 ..

1 MSKSPVAIII LDGFGKRAET VGNAAVAQANK PNFDRYMADE PHGELKAAGL
 51 DVGLEPESQMG NSEVGHNTNG AGRIYQSLT RIDKAIIEGE FQENKALNNA
 101 FTHTKENNSD LHLFGLLSDG GVSHSHNHLY ALLETANDKG VKNVYTHAFL
 151 DGRDVAPOSS LEYLETLOKA ISDLNLYGALA TVSGRFYAMD RDKRMVEREK
 201 AYKAIYSAEG EKFEPIELV KASYANDKND EFTVPAITTK DGRPAVATVKD
 251 NDAVIEFENR PDRAIOLNSA FTDKEMDHD RGADHPKNIK FVTMLTYNPS
 301 IDAEVAFEPI EMKNVIGEV L SNEGSLQRLI AETEKYPHTV FPMNGRNEE
 351 FPGENRLIN SPRKETYDQ PEMKAYEYTD ALVEDINDK HDAILIINFAN
 401 PDWVGHSGML EPTIKAIEN DENIGRYVDL ILEKGSALII FADHGNSETM
 451 SFPEGKPHTA HTTVPVPIV TKKGVTLREG GRLADVAPTM LDLLGVKKPA
 501 EMTGESLIQK

11AA-SEQUENCE 1.0

ID ABB49398 standard; Protein; 289 AA.

AC ABB49398:

DT 05-FEB-2002 (first entry)

DE Listeria monocytoenes protein #2102.

KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KM vitamin B12; bacterial infection; disease.

OS Listeria monocytoenes.

PN WO2001/7335-A2.

PD 18-OCT-2001.

PF 11-APR-2001; 2001WO-FR01118.

PR 11-APR-2000; 2000FR-0004629.

PA (INSP) INST PASTEUR.

PI Buchleser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
 PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst P, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;

PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 DR WPI; 2002-010914/01.

PT Genomic sequence for Listeria monocytoenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and
 PT related polypeptides

PS Claim 6; SEQ ID No 2103; 192pp; French.

CC The present invention relates to the genome sequence of Listeria
 CC monocytoenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytoenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytoenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytoenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by L.
 CC monocytoenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 289 AA:

ABBA9398 Length: 289 May 30, 2002 09:25 Type: P Check: 1607 ..

1 VMKRVVYKVK QNCIVQVQGA VSARVGRNDV SCNAQAOLAY MLFSIRPMLL
 51 IAAITLALYLH IDKDSVFNNI KEFAPDQIMD FLEENLNNLL TORNGGLIST
 101 GIATLMSAS NGMAAVMKSL NKAYGVTKR NYVQRLISM FFLAMLATV
 151 GATILLLVFG QQIGMFLINH LNFSEDFLSF WNNIRWYVTL IVTFVYVTFPL
 201 YWAPNRRST LISVDPALF STIGWTVASV GFAVYVNNFG NYSATYGSIG
 251 VIILMLMFY LFGIILMIGG ELNATLATRK KKEIGEIN

This Page Blank (uspio)

Thu May 30 09:36:22 2002

11AA_SEQUENCE 1.0
 PI:G71532 - probable threonine/tyrosine-specific protein kinase (BC 2.7.1.-) -
 Chlamydia trachomatis (serotype D, strain UW3/Cx)
 C:Species: Chlamydia trachomatis
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: G71532
 R:Stevens, R.S.; Kalman, S.; Lammel, C.J.; Pan, J.; Marathe, R.; Aravind, L.;
 Mitchell, W.P.; Olinger, L.; Tatusov, R.L.; Zhao, Q.; Koonin, E.V.; Davis, R.W.
 Science 282, 754-759, 1998
 A:Title: Genome sequence of an obligate intracellular pathogen of humans:
 Chlamydia trachomatis
 A:Reference number: A71570; MUID:99000809
 A:Accession: G71532
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-934 <ARN>
 A:Cross-references: GB:AE001302; GB:AE001273; NID:g3328708; PIDN:AAC67894.1;
 PID:g3328716
 A:Experimental source: serotype D, strain UW-3/Cx
 C:Genetics:
 A:Gene: pknD
 C:Superfamily: Chlamydia trachomatis probable threonine/tyrosine-specific
 protein kinase; protein kinase homology
 C:Keywords: phosphotransferase
 F:2-293/Domain: protein kinase homology <KIN>
 G71532 Length: 934 May 30, 2002 09:25 Type: P Check: 4900 ..

1 MQRVEIRIRI GKGGMGEYLL AHDKACSRV ALKRIEDLS GNALLQRRL
 51 REAKIADILI HPGIVPVYSI CSDEAVYYT MPYIEGSLK SLKSWQKE
 101 VLKSELEERT SVKSFPIFD KICATVEYIH SKGVLRHDLK PNILLGLRG
 151 EYVITMGMA IEFKAKEKL EGDDEAVSF DENICSSM TIRGRVGRP
 201 DYAPESLIG VEASEKTDYI ALGLITYOML TLAPPYRRKK GRKLSYRDYV
 251 LPIEKSPPYR EIPPSLSOIA MKAIAINPAD RFSIOELRQ ALQPYLQGP
 301 EMTVKTTLMA KEKSCWKYYD PILSRFPV LASSPQWYN FMISEVISA
 351 STRREYTVTK SAVHEGMCIL FLPSKEARG EYCGIGLWF SYONHEITVS
 401 LKNGIEIQR KSOEMISQOS RFALLIEKSD NRIAVEEQA LPIIHIDYLP
 451 SLGNRIQVVI QDLOGMSNTA ISESIGALRV SCLAVPDAFL SEKLYDQAAI
 501 FYRKIRDSEP GRKESYEAOE RLGVTLTLQI EEOGGDLTQA LSEPDYLGCG
 551 AGAPLEYLCK ALVYORNGSF VEETIRCLLFA LKRYSQHEI PRLEDHLCFR
 601 LYSLSLKHNS EALVFMILLI WIAPEKISVR EEEERLRIIT HGOQALFECQ
 651 VDKAPLQFNS SMELFLSFW TGFSLELPDL FRRAGGLRDY QALADIFYVA
 701 GVSNGREAFM OFSTALANYS DETEPESLH NQKVAELMF YGVEALRNK
 751 DYQKAKKLM KTFPLQLLYA LDMFHQAFL DELESTIDL LQAIYPASE
 801 EERDHLIYVI IQTHLMNRDL ERAYKLLNDR FPDDELAET SEAFILMGCV
 851 LALTGRVVVY KAHFSRCRYK YGKSALIGKC VDGDIEDYLD NLVWMEKKMT
 901 LFOSYFLRC LNESPRRYEK YRQAVLSMEN NFPD

11AA_SEQUENCE 1.0
 PI:W5WIR1 - E5 protein - thesus papillomavirus (type 1)
 C:Species: thesus papillomavirus
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 27-Jan-1995
 C:Accession: F38503
 R:Ostrow, R.S.; Labresh, K.V.; Faras, A.J.
 Virology 181, 424-429, 1991

A:Title: Characterization of the complete RhpV 1 genomic sequence and an
 integration locus from a metastatic tumor.
 A:Reference number: A38503; MUID:91135018
 A:Accession: F38503
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-157 <OST>
 A:Cross-references: EMBL:M37717
 C:Superfamily: thesus papillomavirus E5 protein
 C:Keywords: early protein

W5WIR1 Length: 157 May 30, 2002 09:25 Type: P Check: 8739 ..

1 IYVIGTQWS HRPVHTLNS IQVLCKANCC CYACKPPFC CFWLCPCCCF
 51 CLALCFVHLL SRCFCFPPVC LSVAAVAVYL GVHSEPCSF MSVFVLFNP
 101 VAFDFPACPO CGLOQNDVNF AHRVITISYF AIYAVNIYFV LALLVGAARK
 151 ATSRART

11AA_SEQUENCE 1.0
 PI:C69762 - di-tripeptide ABC transporter (membrane pr) homolog yclF - Bacillus
 subtilis
 C:Species: Bacillus subtilis
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: C69762
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azavedo,
 V.; Bertero, M.G.; Bessieres, P.; Bolyen, A.; Borchert, S.; Boriss, R.;
 Boursier, L.; Brans, A.; Braun, M.; Brignelli, S.C.; Bron, S.; Broillet, S.;
 Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani,
 J.J.; Conerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denizot, F.; Devigne, K.M.;
 Duesterhoeft, A.; Ehrlich, S.D.; Emerson, P.T.; Enliant, K.D.; Errington, J.;
 Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi,
 A.; Gallier, N.; Ghim, S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi,
 G.; Guiseppl, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood, C.R.; Henaut, A.;
 Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M.; Jones, L.;
 Joris, B.; Karamata, D.; Kasahara, Y.; Klier-Blanchard, M.; Klein, C.;
 Kobayashi, Y.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.;
 Lapidus, A.; Lardinois, S.
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda,
 S.; Maueel, C.; Medigue, C.; Medina, N.; Mellado, B.; Mizuno, M.; Moestl, D.;
 Nakai, S.; Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogilvara, A.; Oudega,
 B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Porwollik, S.; Prescott,
 A.M.; Presseau, E.; Pujic, P.; Purnelle, B.; Rapoport, G.; Rey, M.; Reynolds,
 S.; Rieger, M.; Rivoita, C.; Rochna, E.; Roche, B.; Rose, M.; Sedate, Y.; Sato,
 T.; Scanlon, E.
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska,
 A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo, B.; Sorokin, A.; Tacconelli, E.;
 Takagi, T.; Takahashi, H.; Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka,
 T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vandenboj, M.;
 Vannier, F.; Vassartelli, A.; Viari, A.; Wandut, R.; Wedler, E.; Wedler, H.;
 Wetzenecker, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,
 K.; Yata, K.; Yoshida, K.
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis.
 A:Reference number: A69580; MUID:98044033
 A:Accession: C69762
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-492 <KUN>
 A:Cross-references: GB:299106; GB:AL009126; NID:g2632653; PIDN:CAB12175.1;
 PID:g2632668
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yclF
 C:Superfamily: peptide transporter protein

C69762 Length: 492 May 30, 2002 09:25 Type: P Check: 9486 ..

1 MASIDNESII KSVQKGFRC HPRGLFTLFF TEFWERESYI GMRALILLYTL
 51 YTTTAVGCGG EFKGTAVALM STYSLYVMS TTIIGWLADR VEGTANFVRY
 101 GGIFIFMGHI ALAYPOSSIA FYISMVLIIV GTGLLKPNVS SAVGDLYTRKE
 151 DRRDSGFSI FYMGINLGL LAPLIVTIG QKYNHILGFG AAVAGMLLGL
 201 IYVPLTRKKN LGLAGSNPN PLSKSAIGT GIGVYIYA VIISVOTGVY
 251 TTKRFIDVS ILGLIPVYI FIIMFTRSKA DKTEKSLAA VYPLFGAVM
 301 FMAIQOGAT ILAVYADERI RLSIGFELQ SSMFQSLNPL FVYIFAPITA
 351 WLMKIKGRQ PSTPYKFSIG IILAGLSFTI MFRPAMQKE ALVSPMLVL
 401 SFLLVVLGEL CLSPVGLSVT TKLAPAFSA QINSMWELTN AAAQAINAOV
 451 AGLFDKIPET MYFTIGLIS IVLGILLTL SPVTKRAMKG VL

!!AA_SEQUENCE 1.0
 PI:C69291 - pheromone shutdown protein (trab) homolog - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: C69291
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, R.J.; Gwinn, M.; Hickey, E.K.; Peterson, J.D.; Richardson, D.L.; Kerlavage, A.R.; Graham, D.E.; Kyrpides, N.C.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Dougherty, B.A.; McKenny, K.; Adams, M.D.; Loftus, B.; Peterson, S.; Reich, C.I.; McNeil, L.K.; Badger, J.H.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Uitterback, T.; Cotton, M.D.; Springs, T.; Artlach, P.; Kaine, B.P.; Sykes, S.M.; Sadow, P.W.; D'Andrea, K.P.; Bowman, C.; Fujii, C.; Garland, S.A.; Mason, T.M.; Olsen, G.J.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon Archaeoglobus fulgidus.
 A:Reference number: A69250; MUID:98049343
 A:Accession: C69291
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-396 <KLE>
 A:Cross-references: GB:AE001082; GB:AE000782; NID:g2689405; PIDN:AAB90915.1; PID:g2650316; TIGR:AF0331
 C:Superfamily: pheromone shutdown protein

C69291 Length: 396 May 30, 2002 09:25 Type: P Check: 4349 ..

1 MEKRLIVG TAHVSKRSVE EVAEIEREK PDAVAVELCP RRYHALVHGO
 51 REEISADVYI RKNVFMILF OLIIAYFQK VGEETGVKRG SEMLAIEKA
 101 RENAGADVLLI DRDGLTFETR FWOKLFEVER IKLIFHLVRS TFSGDEIEVD
 151 EMLEEDVDM LVKEFRKISP NAAKVLDER DVMANMLLN ALSRYNRIYA
 201 VVGAGHRKRI EELALKKEN PVNIRELEV KKRSTFKTF MCAFALLIVG
 251 TFLIATSLN TEVLQAFELY WFLINGVLSA VGASLARGHP ISIVATFLCA
 301 WMTSLNPLVA AGWVGLVEA WIRKPTVEDF SKLVEASLR EMQKQKFRVY
 351 LLYAALTNGV SMIGTIYGVW YIISTFGVDV AKVYGRVAYE VIGGILL

!!AA_SEQUENCE 1.0
 PI:G86502 - s/r protein kinase [Imported] - Chlamydia pneumoniae (strain J138)
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: G86502
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shida, T.; Ishii, K.; Hattori, M.; Kuhara, S.; Nakazawa, T.

Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349
 A:Accession: G86502
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-918 <STO>
 A:Cross-references: GB:BA000008; NID:g9878468; PIDN:BA098305.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: CP10095
 C:Superfamily: Chlamydia trachomatis probable threonine/tyrosine-specific protein kinase; protein kinase homology

G86502 Length: 918 May 30, 2002 09:25 Type: P Check: 36 ..

1 MGEVYLAIDP VCSRKVALKK IREDIAENPL LKRRLFEAR IADLIHPGV
 51 VPYITISEK DRYITMPYI EGYTLKITLK SWQKESLSE ELAEKTSVGA
 101 ELSIFPKICC TIEVYHSRGI LHRDLKPDNI LGLFSEAVI LMGAAVAGC
 151 EEBDLIDIV SKEVLSRM TIPGRIVGTP DYMAPERLIG HPASKSTDYI
 201 ALGVLYQML TISFPYRRKK GKTIYDQOR IPSQEVAPY REIPEPLSAV
 251 VMRLAVDPQ ERYSSVTEIK EDISHLKGK PKWTLTALP PKSSSMKLN
 301 EPILSKYEP MLEVSPASWY SLAISINISF SEMRLEVTLK KGLNCGFI
 351 LFTSENALG GDFYGYGFW LHKERITLV SYKNSLEIQ RQSDLESK
 401 EYFLIALECH NLSLSEVDG TTWLIHMYL PSSSGRAVLI VDMEDILED
 451 IGIFESSGSL RVSCLAVPDA FLAEKLYDRA LVLYRIAES FGKRGYBA
 501 RFRAGITVLE KASTNNDE FALAIERSK LHDGVAPLE YLGALEYOR
 551 LQRYNEIKS LILAIKRSQ HPEIFRLKH VYRLHESFY KDRILAVEM
 601 ILVLEIAPQA ITPGOBEKIL VWLKDKSRAI LFCLLDPTVL ELRSSMELF
 651 LSWWSGFTH LNSLFHRAMD QSDVRALIEI FYAACDLHKM QLSLSDIDF
 701 KESLEOKAT BEIVSEFED LGAFLEAIQS IENKEDAEKI FVSNDQLSPT
 751 LVVYIFDLPA NRALLESQGE AIFQALDLIR SKYPENFYHD YLRNHEIRAH
 801 IWCNREKALS TIFENYTERQ LKDEQHLEFY LYGCYALAIQ GADAKQHED
 851 VCRDRIIFRA SILARNYNRL GLPKDALSYQ ERRLLLRQKF LYFHLGNHD
 901 ERDLQOTMYH ILTEEFQI

!!AA_SEQUENCE 1.0
 PI:C72120 - s/t protein kinase - Chlamydia pneumoniae (strain CWL029)
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
 C:Accession: C72120
 R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.; Davis, R.W.; Stephens, R.S.
 Nature Genet. 21, 385-389, 1999
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A:Reference number: A72000; MUID:99206606
 A:Accession: C72120
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-918 <ARN>
 A:Cross-references: GB:AE001595; GB:AE001363; NID:g4376353; PIDN:AAD18248.1; PID:g4376355
 A:Experimental source: strain CWL029
 C:Genetics:
 A:Gene: CPn0095

C:Superfamily: Chlamydia trachomatis probable threonine/tyrosine-specific protein kinase; protein kinase homology

C71210 Length: 918 May 30, 2002 09:25 Type: P Check: 36 ..

```

1 MGEVYLAADP VCSRRKVALK IREDLAENPL LKRRLREAR IADLHHPGV
51 VPVYTYSEK DPVYTYMPYI EGYTLTKTLK SWOKESLSK ELAEKTSVGA
101 FLSTFHKICC TIEYVHSRGI LHRDLKPDNI LLGLESEAVI LDMGAAVAGC
151 EEDBLDIDV SKEEVLSNRN TTPGRIVGTP DYMADERLLG HPASKSTDIY
201 ALGVLYQML TLESPYRRKK GKRIYLDGOR IPSPQEVAPY REIPFELSAV
251 VMRLAVADPQ ERYSSVTEK EDIESHLKGS PKWTLTTMLP PKKSSMKLN
301 EPILSKYRP MLEVSPASWY SLAISNIEF SEMRLEYTLS KKGINEGFGI
351 LLPTSENALG GDFYQGYGFV LHIKERTLSV SLVKNSELIQ RCSODLESDEK
401 ETFLALEQH NLSLSLEVDG TFWLHMNTLV PSRGRVAII VRDMEDLIED
451 IGIESSGSL RVSCLAVPDA FLAEKLYDRA LVLYRRIAES PPKRKEGYEA
501 RFRAGITVLE KASTDNNQOE PALAIEEFSK LHDGVAADPLE YLAKALVYQR
551 LQENNEEKS LLLALKRYSQ HPEIFRLKDH VYRIRHESFY KRDLALVEM
601 ILVLEIAPQA ITPGQEEKIL VMKDKSRAT LECLIDPTVL ELRNSKMELF
651 LSYMSGFIPH LNSLFRHAWD OSDVRALIEI FYVACDLHKW QFLSSCIDIF
701 KESLEDQAT EEIYERSFED LGAFLEAIOS IFNKEDAKKI FVSNDOQSPI
751 LVYTFIDLEA NRALLSQCE AIFQALDIR SKVEPNFYHD YLRNHEIRAH
801 LMCNRKALS TIFENYTEKO LKDEQHELFV LYCYLALIQ GAEMAKOHFD
851 VCRDRITFPA SLARNNRNL GLPKDALSQY ERRLLLRQKF LYFICLGNHD
901 ERDLQQWYH LITREPOL

```

11AA_SEQUENCE 1.0
 P1:AF1811 - dTDP-glucose 4-6-dehydratase [imported] - *Anabaena* sp. (strain PCC 7120)
 C:Species: *Anabaena* sp.
 A:Note: *Anabaena* sp. (strain PCC 7120) is a synonym of *Nostoc* sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
 C:Accession: AF1811
 R:Kaneko, T.; Nakamura, Y.; Molk, C. P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Irituguhi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsunoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. strain PCC 7120.
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AF1811
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-356 <KUR>
 A:Cross-references: GB:BA000019: PIDN:BA077562.1; PID:g17135016; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: rfbB
 C:Superfamily: *Escherichia coli* UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

AF1811 Length: 356 May 30, 2002 09:25 Type: P Check: 1483 ..

1 MIONDTLELT TSVAAMPILI TGAGAGTGSN FVHHWEKYP GDRMTIVDAL

```

51 TYAGNRQMLA DIEGKANLRF VKGDIDGRAL IDOLLEEKI QAIAPFAES
101 HYDRSIVADP AFIOINVGT FILLFAFRHH WTKGKRPANY RFLVNSIDV
151 YGSELEDDPA FTEPTTYPAPN SPYSASKAGS DHLARAYYHT YGLPPLITNC
201 SNNGPYRFP EKLIPLICLN ILGKRLPIY GDGLNIRDLV YVEDHCRALD
251 IYHGRKGE TYNGGNNEI KNIDLYOMIC ELMDLADL PVSASKLIT
301 FVKDRPGHDR RYAINATKIK TELGWEPQOT ISTGLRHTIQ WYLTNRHWE
351 ALPKKE

```

11AA_SEQUENCE 1.0
 P1:AH1381 - phosphoglycerate mutase homolog pgm [imported] - *Listeria monocytogenes* (strain EGD-e)
 C:Species: *Listeria monocytogenes*
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 C:Accession: AH1381
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, H.; Brandt, P.; Chakraborty, T.; Chablit, A.; Chetoui, F.; Couve, E.; de Daruvar, A.; Deloux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kretz, J.; Kunz, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitounam, A.; Mata Vicente, J.; Ng, E.; Nordstiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland, J.; Cossart, P.
 A:Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AH1381
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-510 <GLA>
 A:Cross-references: GB:NC_003210: PIDN:CAD00534.1; PID:g16411944; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: pgm
 C:Superfamily: phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent

AH1381 Length: 510 May 30, 2002 09:25 Type: P Check: 2157 ..

```

1 MSKSPVAIIT LDGFGKRAET VGNAVAQANK PNFDRYWADF PHGELKAAGL
51 DVGLPEQGMG NSEYGHNTIG AGRIVQSLT RIDKALIEGE FOENKALNNA
101 FTHRKENSND LHLFGLLSDG GYVSHINHLV ALLEFAKDG VKNYITIAFL
151 DGRDVAPOSS LEVLETLQKA ISDLNAGAIA TVSGRFYAMD RDKMERVEK
201 AYKALVSAEG EKFEDEIELV KASYANDKND EFVVAITTK DGKPVATYKD
251 NDAVIEFNR PDRAIQUSNA FTDKEMDHD RGADHPKNIK FYVTITLNP
301 IDAEVAFEPi EMKNVIGEV L SNEGLSOLRI AETERYPHV FFMNGRNEE
351 FPGENRIILIN SPKVETYDLQ PEKSAVEYND ALVEDIKNDK HDALITLNPAN
401 PDMMGHSGML EPTIKALEAV DENLGRVVDL ILEKGSALII FADHGNSETM
451 STPEGRKPTA HTTVVPVPIV TKKGVTLRBG GRLADVAPTM LDLLGVKPPA
501 EMTGESLIQK

```

11AA_SEQUENCE 1.0
 P1:AT1750 - phosphoglycerate mutase homolog pgm [imported] - *Listeria innocua* (strain C11p11262)
 C:Species: *Listeria innocua*
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: A11750
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioeker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussuguet, O.; Eitlin, K.D.; Fsihl, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hahn, T.; Hauf, J.; Jackson, D.; Jones, I.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Mañeno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablo, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund, J.; Cossart, P.
 A:Title: Comparative genomics of *Listeria* species
 A:Reference number: ABI077; M01D:21537279; PMID:11679669
 A:Accession: A11750
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-510 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CAC97777.1; PID:q16415072; GSPDB:GN00178
 A:Experimental source: strain C11p11262
 C:genetics:
 A:gene: pgm
 C:superfamily: phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent
 A11750 Length: 510 May 30, 2002 09:25 Type: P Check: 2291 ..
 1 MSKSPVAIII LDGFGKRAET VGNAVAQANK PNFDRYWANF PHGELKAGCL
 51 DVGLEPGQMG NSEVGHNTIG AGRIYQSLT RIDKAIEGE FOENKALNNA
 101 FHTTKENNSD LHLFGLSDG GVSHINILV ALLETADKG VKNVYTHAF
 151 DGDVAPQSS LEYLETLEKA ISDLNYGAIA TVSGREYAMD RDKMERVERK
 201 AYKAIVSAEG EKFEDEPIELV KASYANDKND EFVVAITTK DGPVATVMD
 251 NDVIVFNFR PDRAIQLSNA FTDKEMHFD RGHHPKNIK FYTMILYNSS
 301 IDAEVAFEP1 EKNVIGEVY SNEGISOIRI AETEKYPHT FPMNGRNEE
 351 FPGENILIN SPKVEYDLK PEMSAEYVD ALVEDIKND HDAILNFAN
 401 PDVGHSGML EPTIKAIKAV DENIGRAVDL ILEKGSATIT PADHGSSEIM
 451 STPEGRPHRA HTTVPPVPIV TKKGVTLEEG GRLADVAPIM LDLGVKKPA
 501 EMTGESLIQK
 I:AA_SEQUENCE 1.0
 F1:A27671 - spectrin alpha chain, nonerythroid - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 29-Sep-1999
 C:Accession: A27671
 R:leto, T.L.; Fortugno-Erikson, D.; Barton, D.; Yang-Feng, T.L.; Francke, U.; Harris, A.S.; Morrow, J.S.; Marchesi, V.T.; Benz Jr., E.J.
 Mol. Cell. Biol. 8, 1-9, 1988
 A:Title: Comparison of nonerythroid alpha-spectrin genes reveals strict homology among diverse species.
 A:Reference number: A93097; M01D:88094373
 A:Accession: A27671
 A:Molecule type: mRNA
 A:Residues: 1-475 <LEI>
 A:Cross-references: GB:M19726; NID:9205641; PIDN:AAA1678.1; PID:9205642
 A:Note: The authors translated the codon GGC for residue 18 as Ser, GAG for residue 45 as Ala, CCG for residue 46 as His, CAG for residue 48 as Pro, AGT for residue 56 as Thr, ATC for residue 61 as Ser, ACA for residue 68 as Gln, CCG for residue 75 as Leu, and GAG for residue 76 as Ala
 C:Superfamily: spectrin alpha chain; calmodulin repeat homology; SH3 homology; spectrin/dystrophin repeat homology
 C:Keywords: actin binding; EF hand
 F:205-311/Domain: spectrin/dystrophin repeat homology <SP18>
 F:320-425/Domain: spectrin/dystrophin repeat homology <SP19>

A27671 Length: 475 May 30, 2002 09:26 Type: P Check: 7377 ..
 1 RMDDEESWT KKKLLVSGE DYGRDLTVQ NLKRNKRL GRLAEQDAI
 51 OGVLDSGKLL IDNTIGREE IQORPEQVE HMKELQGLAA ARQGRLEESL
 101 EYQFPAVNE EEEAMINEMA TLVASEDYGD TLAIQGLK KHPAFETDPT
 151 VHKDRVNDV TNGODLIKN NHHEENISSK MKGNGKVSQ LKXMAQRKA
 201 KLDNSAFIQ FMWKADYVES WIGEKENSIL TDYGRDLSS VQTLTKQET
 251 FDAGLQAFQO EGIANITALK DQILAARIQ SKALEAHAS LKRWITQLA
 301 NSATRRKKLL EAOSHERKVE DLFLAFAKA SAFNSWEMA EEDLTPVRC
 351 NSLEIKALR EAHDAFRSSL SSAQADFQQL AELDRQKSF RVASNPYTF
 401 TMEALEETWR NLQIKERE LEIQEQRRQ EENDKLQEF AOHANAQHW
 451 IOETRYILD GSCIGRRVGN SGISA
 I:AA_SEQUENCE 1.0
 P1:T32253 - hypothetical protein T15B7.16 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T32253
 R:Pauley, A.; Gattlung, S.
 Submitted to the EMBL Data Library, September 1997
 A:Description: The sequence of *C. elegans* cosmid T15B7.
 A:Reference number: Z21139
 A:Accession: T32253
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-491 <PAU>
 A:Cross-references: EMBL:AF022985; PIDN:AB69969.1; GSPDB:GN00023; CESP:T15B7.16
 A:Experimental source: strain Bristol N2; clone T15B7
 C:genetics:
 A:gene: CESP:T15B7.16
 A:Map position: 5
 A:Intons: 25/2; 110/3; 172/3; 304/3; 330/2; 370/2; 403/3; 432/1
 C:superfamily: acetylcholine receptor
 T32253 Length: 491 May 30, 2002 09:26 Type: P Check: 8670 ..
 1 MEYFKQNLKK KLTQNSPRE HSFGEPPAND TISQVLRL FDDNYKNKTY
 51 PLKTSATVVS VEEGIONIAQ VSEISASFTL DLEFSQIWHQ PRIRPHILTN
 101 CLQNLTLGIS MVEKLTMPNV CFVNSKTEI HSSETPNIFL MIYPNGTVMV
 151 NYRLQVSPC MYDLVLEPMQ IMNCETLES YAYNAKVKL NMRMQPVES
 201 IAKSKLSDFT LVGLQWTKNS FEYAGQMDQ VSSPSCINCF SKMDGQQLTV
 251 SLTFSRAYGF YLLQMTIPTV SSVLEFSVSF WIDIKALPAR ITLVGSSILMA
 301 LTFQYGNVAK NLPVGYVKS IDVYVLTFTA FIFLTNIEVA FQCYLDEENN
 351 LRRKQRAEK KKEKRAVILOR KDKRKNNGV ATYTNNMAA ESLRBNYDE
 401 PFSQNGTTSK KISISQANN MFESLHALAQ FGLITDDDE NTKWTAQNVQ
 451 KFCRAKAPLS FCFNLNIYMC YLYQNYVLAK AEALAQTPP S
 I:AA_SEQUENCE 1.0
 P1:S40462 - t-complex-type molecular chaperone top1 (clone ASTCP-K36) - oat
 C:Species: *Avena sativa* (oat)
 C:Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change 20-Aug-1999
 C:Accession: S40462; S39320
 R:Edmann, B.; Krenz, M.; Mummert, E.; Schaefer, E.
 FEBS Lett. 336, 313-316, 1993
 A:Title: Two top-1-related but highly divergent gene families exist in oat

encoding proteins of assumed chaperone function.

A:Reference number: S40461; MUID:94085629

A:Accession: S40462

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-535 <EIM>

A:Cross-references: EMBL:X75778; NID:9435174; PIDN:CAA53397.1; PID:9435175

A:Note: the nucleotide sequence was submitted to the EMBL Data Library,

November 1993

C:Genetics:

A:Gene: top1

C:Superfamily: molecular chaperone t-complex-type

C:Keywords: molecular chaperone

S40462 Length: 535 May 30, 2002 09:26 Type: P Check: 4989 ..

```
1 MALPDFEYR PFILREOEK KSRLOGLDNQ KANIAGASV ARIITSLGP
51 KGMKMLQSP DGDVITINDG ATILEMDVD NOIAKIMVEL SRSODYDIG
101 GTTGVMVWAG SLLEQAELK EGIPIPIVA EGYEMASRLA VDHEISISTK
151 YEFATDIEP LVOTCKTTLS SKIVSRCKRA LAETIVKAVL AVADLERKDV
201 NLDLIKVEGK VGSKLEDTEL VOGIIVDKDM SHPQPKRIE DAHIAIITCP
251 FEPPKPKTKH KVIDIVVEKF QTLRGQEQKY FDEMVOCKCKD VGATLVICQW
301 GFDEBANHL MORELPAVRW VSGVELELIA IATGRIPIPR FOELSTERLG
351 KAGIVREKSF GTTKDRMLYI EKCANSKAVT IFIRGNKRM IETKRSIHD
401 ALCVARNLII NSIIVGGGS AEISCSIAVE AAADRHGVE QYAIRAFADA
451 LDAIPLALAE NSGIPIIDTL TVVKSQHYKE NNSRCIGICN DVGINDMKEQ
501 NVFETLIGKQ QOILLATQV KMLIKIDYI TPSEY
```

IIAA-SEQUENCE 1.0

P1:PC2306 - dnaJ protein - *Synechococcus* sp. (strain PCC 7942) (fragment)

C:Species: *Synechococcus* sp.

C>Date: 15-Feb-1995 #sequence_revision 26-May-1995 #text_change 20-Jun-2000

C:Accession: PC2306

R:Nimura, K.; Yoshikawa, H.; Takahashi, H.

Biochem. Biophys. Res. Commun. 201, 848-854, 1994

A:Title: Sequence analysis of the third dnaK homolog gene in *Synechococcus* sp.

PCC7942.

A:Reference number: PC2306; MUID:94271242

A:Accession: PC2306

A:Molecule type: DNA

A:Residues: 1-189 <NIM>

A:Cross-references: DDBJ:D29968; NID:9559383; PIDN:BAA06235.1; PID:d1006799;

C:Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology

F:6-71/Domain: dnaJ amino-terminal homology <DNJ>

PC2306 Length: 189 May 30, 2002 09:26 Type: P Check: 7087 ..

```
1 MONEFDYAL LGIPQADQA AIKAARFRLA ROCHPDLNG DROAEERKQ
51 ISEAVEIISD PDRAEYQRF SRYWQOOGAA SVGSDDYGD FPPDIFPVE
101 LLGRRTVERS PRSARSAAS TSSALSRLDE RSLVDPKTA LQGSQAQLQ
151 EDGRLEVDI PAGIQAGEYL RLNGGKIGG DLLRVQLQ
```

IIAA-SEQUENCE 1.0

P1:JC5550 - dnaJ protein - *Synechococcus* sp.

C:Species: *Synechococcus* sp.

C>Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 20-Jun-2000

C:Accession: JC5550

R:Oguchi, K.; Nimura, K.; Yoshikawa, H.; Takahashi, H.

Biochem. Biophys. Res. Commun. 236, 461-466, 1997

A:Title: Sequence and analysis of a dnaJ homologue gene in cyanobacterium *Synechococcus* sp. PCC7942.

A:Reference number: JC5550; MUID:97382457

A:Accession: JC5550

A:Molecule type: DNA

A:Residues: 1-287 <OCUD>

A:Cross-references: DDBJ:AB003519; NID:g2293346; PIDN:BAA21679.1; PID:g2293347

A:Experimental source: strain PCC7942

C:Comment: This protein cooperates with DnaK protein in protein folding under various stress conditions as well as under normal conditions. In the mechanism of action of the DnaK chaperone machine, this DnaJ stimulates the hydrolysis of DnaK-bound ATP.

C:Genetics:

A:Gene: dnaJ

C:Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology

F:6-71/Domain: dnaJ amino-terminal homology <DNJ>

JC5550 Length: 287 May 30, 2002 09:26 Type: P Check: 2256 ..

```
1 MONEFDYAL LGIPQADQA AIKAARFRLA ROCHPDLNG DROAEERKQ
51 ISEAVEIISD PDRAEYQRF SRYWQOOGAA SVGSDDYGD FPPDIFPVE
101 LLGRRTVERS PRSARSAAS TSSALSRLDE RSLVDPKTA LQGSQAQLQ
151 EDGRLEVDI PAGIQAGEYL RLNGGKIGG DLLRVQLQ SNFOVQGSVD
201 IYTLNVSAM AVLGQVTVR TLDGPVQMKL PASLSRGQRL RLAKGYSKP
251 SGDRGDQIVY IQQLPTRL S PEERQLEYOL RSLQSR
```

IIAA-SEQUENCE 1.0

P1:JC5648 - terminal protein precursor - Ovine adenovirus OAV287

C:Species: Ovine adenovirus OAV287

C>Date: 28-Oct-1997 #sequence_revision 28-Oct-1997 #text_change 05-Nov-1999

C:Accession: JC5648

R:Virati, S.; Brookes, D.E.; Boyle, D.B.; Both, G.W.

Gene 177, 35-41, 1996

A:Title: Nucleotide sequence of ovine adenovirus tripartite leader sequence and homologues of the Iva2, DNA polymerase and terminal proteins.

A:Reference number: JC5648; MUID:97080497

A:Accession: JC5648

A:Molecule type: DNA

A:Residues: 1-641 <VRA>

A:Cross-references: GB:U01557; NID:g1117828; PIDN:AAC55958.1; PID:g1117831

C:Superfamily: adenovirus terminal protein

JC5648 Length: 641 May 30, 2002 09:26 Type: P Check: 4120 ..

```
1 MPTAGSNSSF SFLPDVPEL LDKSGVSNDE RRFCKIGCIC SNKINFLPAL
51 QIHQMOQLTQ QSQHTLRRLR LTLDLNNRSI AHSRTIHEG IKMARFENY
101 PVTQLDLNR RGYVTSOPPE EGEPPMLLI GFYVARVLN QYLEDQRTYS
151 NISYKLYLSP ISFERMTWQ ILTDCSYSIN TGSYMAION VENFSQITISQ
201 IONAVIMDKI LGSIQADMQ GFGSAISAQ QNRSFQOHQ PFSQAFWVLN
251 IODRDTYLIK LICKIKKALC KFLILSMHN SECIDLPRS EFWIELTDE
301 FSQLEIPEAE NRINLKLAT VLTGKGGMH GGALTLRSGT RVGLPFRLRP
351 RENYRAITFI MRRSGDVIR RFIDRLPVNR RQVRVVEET SSSPLTIVEI
401 EDEGFSDRPE SSSDLSREEF NDEVIASIVD LIONLEEELT PEARSNFPN
451 YGSEFFQLLI RFYNENRLTD DFIQKWLVEF FILEHVA STL YLYLNLVON
501 RLAAKNIGIQ FVQIILGRN ENGEDITFR WFNREQAQR QLYSITIDE
551 LGIIEANERS YSFSTPEERD QLLQDIDIVE DSGSIEEVIN QVNTDFSDLD
```

601 SVEIAFRIKL SGIVGSTNE VILRSFERVR EALNRMLÖR Q
11AA_SEQUENCE 1.0
P1: T08841 - polyprotein - douroucouli hepatitis GB virus A
C: Species: douroucouli hepatitis GB virus A
C: Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C: Accession: T08841
R: Eker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.;
Kushnitar, I.K.
J. Gen. Virol. 79, 41-45, 1998
A: title: Genomic analysis of two GB virus A variants isolated from captive
monkeys.
A: Reference number: Z16486; MID: 98120818
A: Accession: T08841
A: Status: translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-3005 <ERK>
A: Cross-references: EMBL:AF023425; NID:92828599; PIDN:AAC40502.1; PID:92828600
C: Superfamily: hepatitis C virus genome polyprotein
C: keywords: polyprotein

T08841 Length: 3005 May 30, 2002 09:26 Type: P Check: 8233 ..

1 MEVLLVLLK TALAGAFLOP ASHACHAGY YLLNCCSDS EITACPDGC
51 LVSVGCTVCD GRCDLIRPG VATRGHPGG ELFGALMSTS SPVVAAYVA
101 GITLGEPPS AALALALAR VEMTPRPNT CLMDCDLALQ SEAMDLVEDM
151 RDTFWYKMY MTLPMRLWYG LVGMSFALMV VVLLLEQR VVAVLLIAM
201 AGSVEPTGG CRCHNGSIV AIPGAKPEDR PNATQVCYC FGOYMLPAL
251 CAGIARNNGD VSGTRDLDP RCPESMGRG SVMCWGSAH WSWRLATEVR
301 LMEOLPESAL CHEFAVGTD RPHPATVLS THGIPCASC VDRASMCN
351 CTDDCWKTKG NKRLSPEAG LGSRLTALM AHVVDGTES KIYTDGERP
401 RYSSHGCVGT HHTAVVAARN YTVSDIGYV HAIACPNPPS PSALPKLIPG
451 RPNACLTIC KEGRLNTAM QAPGXFAP I ETECNMPRS GVAVCEGYAF
501 DFGSKTGF IBSRAGQOIA RAGSMHPRW LITDYLETLL VMKLAEARL
551 VPLVAVALYV WENNAEAT LRIILPVNI TPAAPMTWA LPVPIVVCPT
601 KAGAVGDALY MAACLAGNAT LGVSSVWG AYGAAEAGAR GLRWMGGR
651 SFGAGLAWLT NVGAYLPVE AAVGEFVSA PLVYMAEDS IAVTILMAN
701 IMAVSDHLG VKIAGLVAAH LAKGALPLVL LVAAXVTRHR HSYLGEVCV
751 SLDDASPSD WSMWALAAYV SMCLLTLLGL THGRLRLKLE FYSYWCRFHQ
801 AVRMRLCSK VGRGGRDWRV TAVVVAAGII FPREVVCSA ALTALALLD
851 SIDYLETLL LTAQPARAA RLDSLTFLG DADLTFRAVR RLERRGVLE
901 QHGGVSKXA AAILXDLGA LEPVSTARD CYIYRDART LACGQWVEGL
951 PVVARRGDEV LVGFPPSVRA LPPGFVPTAP VVWQORGLG FSVVKSMLG
1001 RDEREHGSI VVLGTSTRS MGTGVNGVY TTFHGSNART LAGPVGPNVC
1051 RMMSPSDVA YPLPSGASC LEPCKGTQS VMCIRNDGAL CIGRLSKLVE
1101 LDIPTETSD RSGSSPILC DEGVVGMV SVLHRGVKT GRYVYPMET
1151 LPRDSOVKSE APPVPGTGF TEAPLYLPTG SKSTRIPLE YTKAGHKVLV
1201 LNSIATVRA MGPYMEKLSG QHPSIYCGHDTTAVSRTGS PLTYCYGRF

1251 MANRRIRLG ADIVICDECH VTDPTSVLGM GRARLLAREC GVRLLLEFATA
1301 TPPEAPLAQH ESIKEVDIAG DGEVAFYGHK LPVERYRTGR HLLFCHSKVE
1351 CNRLHAALST AGCNVAYYR GMEQLPRAD VCVCATDALS TGTGGSTV
1401 TDCGLMVEEV VEVTLDPTIT ISVRTTPAPA ELRAQRGRG GRGSGGTYYY
1451 AMTASAPAGT LRSGLPMAAV EAGVAMYNLE PDMTADILRA YDACPYTAI
1501 TASVGEALNF FSGLVPMRNY PQVANAQSHQ HNWPLLVGYQ RTMQCEANVA
1551 GPDGPPEMAG LAGTGPPLL CRWGAPRPS VAPHHWDDL QARLVAGEGY
1601 SPCYAGPILL VGLALAGAV LAHMTGSLVY VTSWRVNGNG NPLIQSTRG
1651 VTSAPYQLT VCVEGQTRA DCKCAEAVQ LXESTCGNGP MAASFQACM
1701 KVIDSKRPT AAAAVEKSDS LMRSCANNY CPPGGATSA SAEFASIDTK
1751 FAQAMDAIFT NGRSLVLGY AGYGARRNP LGVAAAFIMG MSGHQVHR
1801 LAALLLVGV GTMLGTPSVG LAMSGAYFAG GSITSMLSA IVAVLGWMEC
1851 AXNPAASLTFD FLTGRAELKD MFLVSCXAS PGASVAGVAL GLLMSMKKG
1901 VGEDVNRIL TLPRGSVLP DGEFVKSEPT ERVSTILRKM SLSRWVTLV
1951 ERRELDTETP CSSMLDLID WLVRGRYTG RRLKGNPSV RVLVCGTTPG
2001 WGSWVGEGH IEARCAGCI ITADVEEGL DVHYSSRLC SNYLKGTVPY
2051 SAAGSDAEP EWPAGPALYQ IGVAEMVOLY RKDRTLVA GA SSVYHLIDE
2101 LRRAIRGPPM FVCGVGSME APLQRPPLY RAGSVRPED VRSLSHTLA
2151 LPPPPRPPPP LAMPPPPPP EOVWTEEBEA DLRAARAI EAVNERLPVP
2201 NPEAAQALD ALEEAVALSIL PHVGAILLGD CSCESEFGH FIPEPAVDV
2251 PIGMEYQVG PLRDOARDLG DRLAVLGARL ESLAEAPHEA SLUTERTMG
2301 ELIDTLADVQ ARLDVTCRSD TSGSFEQIS LSDSEPTIV EGGLKEVVR
2351 POPYREKDLI RPEGAKRLV TVRQSCADR SATRAFALSL PIAAVTATLS
2401 FDLTDHTVSD SNGRVLDPLE LLQNAVGLI VACRRSSSV SYEYISGAP
2451 LNTGRHQPAP KTRPIGTHIT ADTKVYVTD PNQGERAAK VTIWGRSRY
2501 DAHYRGVASE VLOAKTIKS PGWYDEAIA KVHRAAGAF GSKVTVGNNMT
2551 TPAARAEVN MLAKITQOE VPFTLVTKRE VFPKTKRPK PRICPPID
2601 FRIAEKILG DEGLVAKGIL GKSYLFQYTP NORVKLAVDL WROKKIPRAL
2651 TYDATCFDSS IDERDAVET EVFAAASPDP DLVRLGSSY AEGPMVSPKG
2701 VPLGVKRCRS SGVLTTSSAN SITCYIKVA AARAVGLVDP DFLIADDCV
2751 TIYEDGEDH ADALRYALGN YGYDCKPKTH ASLDTAESCS SYIAECNVGT
2801 ERYWMLSTD M RKLARAASE YSDPVSSALG TILMYPWHP I VRVLLPHIL
2851 IMAFRGGGT DDLVCEVOG NHYSPPLVL PEVLVSJLHG RCLRYRADST
2901 KTKMEAGAL RLGHTTLAF YRKRAGNVT RLRGKGWG RLARALLMHP
2951 GLKEHPSIK SIPGEMATP YEHHTVWS GEKPPWMDV KCLFLGVCAF
3001 LTFALL

11AA_SEQUENCE 1.0

P1:H89844 - hypothetical protein SA0682 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: H89844
R:Kuroda, M.; Ohba, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Aoki, K.; Nagai, Y.; Lian, J.; Ito, T.; Kanamori, M.; Matsunaru, H.; Maruyama, A.; Murakami, H.; Hosoyama, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Hiraoka, H.; Kohara, S.; Goto, S.; Yabuzaki, J.; Kanehisa, M.; Yamashita, A.; Oshima, K.; Furuya, K.; Yoshino, C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsuka, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: H89844
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-501 <KUR>
A:Cross-references: GB:BA000018; PID:913700618; PIDN:BA041915.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA0682
C:Superfamily: peptide transporter protein
H89844 Length: 501 May 30, 2002 09:26 Type: P Check: 6701 ..
1 MTOGNSGNG IQDIPQTGF GHPRLGVLF FVEFWERSEY YGMALLIFY
51 MYFAVTDNGL GIDKRTAMSI MSVGSLLYM TSIPGWIAD RITGRTGATL
101 LGAVFIIGH ICLSLPFALI GLFTSMFETI IGSGLMKPNI SNIVGRLYPE
151 NDRMDAGVF IFYMSVNMGA LLSPIILOHF VNVKNFHGF LIAVGAALG
201 LVMYVLENRK NIGSVGNKPT NPLTPAEKKK YGLIIGSVYL AIYLIIVIGA
251 LTNLSFNLV SNTVLVGLA LPIIYFTLII RSKDVTDER SRVAFIPLF
301 ILGWVFAIQ EOGSNVLNII GIEHSDMKLN LFGWKTNCE AIPQSNPLF
351 ILLIAPITSL LMOKLGTQKP SLTPKFAIGT FLAGASYILI GIVGYASGSS
401 NFSVNWYLS YIICVIGELC LSPGNSNAV KLAPKAFNAQ MMSIWTYINA
451 SAQAINGTIV KLEIPLGQFN YFIFLGVAI IYTTIVLAFS PLIIKANKGI
501 R
!!AA_SEQUENCE 1.0
P1:C97003 - ABC-type spermidine/putrescine transport system, permease component II CAC0838 [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: C97003
R:Noelling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, H.M.; Dubois, J.; Qiu, D.; Hittl, J.; Wolf, Y.I.; Tatusov, R.L.; Sabathe, F.; Doucette-Stamm, L.; Soucaille, P.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum.
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: C97003
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-260 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK78814.1; PID:915023730; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0838
C:Superfamily: spermidine/putrescine transport system permease protein polt

C97003 Length: 260 May 30, 2002 09:26 Type: P Check: 8409 ..
1 MVEKWLKREY LTLTYFLVA PIVFLMVFSE NSEKFSHWG HFSLTWYKAL
51 LQDRILITAL YTVVLAIVS SIISTIFGTI SAIGSKNSP LPKULLNVN
101 NIPVLPNDIV MAVSLMTLFT FKIPPEGLT LIAHIAESV PYVILSVLPK
151 LTQLPTDIVK ALDLGATPS YAMRKILILPQ IKSIGIAGFL FAFWMSIDDF
201 VISFPNIGND VTNLSIFITS MARGITPEI NALSLTMEVT ILILLILANR
251 KSISKGEKK
!!AA_SEQUENCE 1.0
P1:C84811 - phosphate transporter (AtPPT2) [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: C84811
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carreira, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: C84811
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-534 <STO>
A:Cross-references: GB:AE002093; NID:93928081; PIDN:AAC79607.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g38940
A:Map position: 2
C:Superfamily: probable inorganic phosphate transport protein PHO84
C84811 Length: 534 May 30, 2002 09:26 Type: P Check: 605 ..
1 MAREQLQVYN ALDVAKTQWY HFTAIITAGM GFTDAYDLF CISTVTKLG
51 RIYHVEGAG KPGTLPPNVA AAVNGVAFCC FLAQGLFQW LGDKLGRKKY
101 YGNTLMWVWL CSTASGLSPG HEPRKAVMRL CFFRWLQFG IGGDYPISAT
151 IMSEYANKKT RGAFFSAVFA MGGFGIMAGG IFATIISSAF EAKPEPPAYA
201 DDALGSTIPQ ADLVWRITLM AGAIPAMTY YRSKMPETA RYVALVAKDA
251 KOASDSKRV LQVEIEPEQO KLEIEISKES KAFGLFSKEF MSRGHLHLG
301 TTSTWFLDI AFYSQNLFOK DIFSAIGWIP PAQSNNAIOE VFKIRAOQL
351 IALCTVPQY WFYVAFIDVI GRFAIQMGF FEMTVFMFAL AIPYNNWTHK
401 ENRIGFVIMY SLTFEPANFG PNATTFVVA ELFPARFST CHGISAASGK
451 LGAMVGARGF LYLAQNPDKD KTDAGYPPGI GVRNSLIYLG VVNFGLILFT
501 FLVPSKSKS LEEMSGENED NENSNDSTRT VPIV
!!AA_SEQUENCE 1.0
P1:T07607 - phosphate transport protein 2 - potato
N:Alternate names: inorganic phosphate transporter 2
C:Species: Solanum tuberosum (potato)
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
C:Accession: T07607
R:Leggewie, G.; Willmitzer, L.; Riesmeier, J.W.
Plant Cell 9, 381-392, 1997
A:Title: Two cDNAs from potato are able to complement a phosphate

uptake-deficient yeast mutant: identification of phosphate transporters from higher plants.

A:Reference number: 207579; MUID:97246321
 A:Accession: T07607
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-527 <LEG>
 A:Cross-references: EMBL:X88891; NID:g1420872; PIDN:CAA67396.1; PID:g1420873
 A:Experimental source: cv. Desiree
 C:Comment: Expressed mainly in root organs when plants are deprived of phosphate and to a lesser extend under sulfur deprivation conditions.
 C:Genetics:
 A:Gene: PT2
 C:Superfamily: probable inorganic phosphate transport protein PH084
 C:Keywords: membrane protein; phosphate transport

T07607 Length: 527 May 30, 2002 09:26 Type: P Check: 2263 ..

1 MAVEDNNLQV LNALDLAKTQ LYHFAIIIA GKGFTDAYS LFSISLVTKL
 51 LGRLYYTKPD LKPGTLPPA RVGLRHGAL VGTLAQLFF GCARLAKMGR
 101 KKYVGMVLV MWCVAASGL SLGNTPKYVM TTLCPFFRWL GFGIGGDYPL
 151 SATIMSEYAN KKTGAFIAA VFAMGFGIL FSGIVALITA AGPDHAYKAP
 201 TRBNNAVST VPOADYIWKI ILMGSLDPA LTYWRMKMP ETARYATALVA
 251 KQARRAADM GRVLOYEIES EEAKIEQISR DETNQGLFS WEVRRRHGLH
 301 LFGTGSTWFL LDIAFYSONL FQKDYFSAVG WIRKAPIMNA VQELYYIARA
 351 QTLIALCSTV PGYWFVARI DIIGRFAIOL MGFEFMTVFM FALADYHNH
 401 TLEARNIGFI VMYSLTFEPA NFGPNATTFV VPAEIFFPARL RSTCHGISAA
 451 AGRAGAIVGA YGELYAAQSK DPMKTDAGYP AGIGIKNSLI VLGFINALGM
 501 VCTFCVPESK GKSLEBASOE TISTGEA

!!AA_SEQUENCE 1.0
 P1:T07604 - phosphate transport protein PT1 - potato
 N:Alternate names: inorganic phosphate transporter 1
 C:Species: Solanum tuberosum (potato)
 C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
 C:Accession: T07604
 R:Leggiewie, G.; Willmitzer, L.; Riesmeier, J.W.
 Plant Cell 9, 381-392, 1997
 A:Title: Two cDNAs from potato are able to complement a phosphate uptake-deficient yeast mutant: Identification of phosphate transporters from higher plants.
 A:Reference number: 207579; MUID:97246321
 A:Accession: T07604
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-540 <LEG>
 A:Cross-references: EMBL:X98890; NID:g1420870; PIDN:CAA67395.1; PID:g1420871
 A:Experimental source: cv. Desiree; root
 C:Genetics:
 A:Gene: PT1
 A:Note: expressed in roots, tubers and source leaves as well as in floral organs
 C:Superfamily: probable inorganic phosphate transport protein PH084
 C:Keywords: phosphate transport

T07604 Length: 540 May 30, 2002 09:26 Type: P Check: 8926 ..

1 MANDLQVINA LDVAKTOLYH FTAIVAGMG FETDAVDLFC ISMYTKLGR
 51 IYHHNNAIAK PGSLPPNVA AVNGVAFPGT LAQOLFPGWL GDMGRRKYV
 101 GMTLMIVIC SIASGLSPGH TPKSVMTTLC FFRFWLGFGL GGDYPLSATI
 151 MSEYANKKTR GAFIAAVFAM QGFGILAGGM VALIYSAFK GAFPAPIYEV

201 DALASTVSOA DEFWRIILMF GAIPAGLTYW WRMKPEPTAR YTALVAKNLK
 251 QAANDMSKVL QVEIEAPEK VAISVANGA NEFGIFSKEF LRRHGLHLIG
 301 TASTWFLDDI AFYSQNLFOK DIFSALGWIP PAQTMNALLEE VYIARAQTL
 351 IALCSTVPGY WFTVAFDRI GRFAIQLMGF FEMTVFMFAL ALPHYHWTLK
 401 DNRIGFVVMY SLTFEPANFG PNATTFVPA EIFPARLRST CHGISAAAG
 451 AGAMVGAFGF LYAAQPTDCK KTDAGYPPGI WCEELVDRPW LCNFLGMLFT
 501 FLVPESKGS LEMSRENEG EETVAEMRA TSGRTVLFEK

!!AA_SEQUENCE 1.0
 P1:T07740 - probable inorganic phosphate transport protein 1 - tomato
 C:Species: Lycopersicon esculentum (tomato)
 C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
 C:Accession: T07740
 R:Daram, P.; Brunner, S.; Persson, B.L.; Amrhein, N.; Bucher, M.
 Planta 206, 225-233, 1998
 A:Title: Functional analysis and cell-specific expression of a phosphate transporter from tomato.
 A:Reference number: 216107; MUID:98408214
 A:Accession: T07740
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-538 <DAR>
 A:Cross-references: EMBL:Y14214; NID:g2266671; PIDN:CAA74607.1; PID:g2266672
 A:Experimental source: cultivar MoneyMaker; root hair
 C:Genetics:
 A:Gene: LePT1
 C:Superfamily: probable inorganic phosphate transport protein PH084
 C:Keywords: transmembrane protein

T07740 Length: 538 May 30, 2002 09:26 Type: P Check: 4193 ..

1 MANDLQVINA LDVAKTOLYH FTAIVAGMG FETDAVDLFC ISMYTKLGR
 51 LYHHNNAIAK PGSLPPNVA AVNGVAFPGT LAQOLFPGWL GDMGRRKYV
 101 GMTLMIVIC SIASGLSPGH TPKSVMTTLC FFRFWLGFGL GGDYPLSATI
 151 MSEYANKKTR GAFIAAVFAM QGFGILAGGM VALIYSAFK GAPPADAYEV
 201 DAIGSTVPOA DEFWRIILMF GAIPAGLTYW WRMKPEPTAR YTALVAKNLK
 251 QAANDMSKVL QVEIEAPEK VTAISEAKGA NDEGLFTKEF LRRHGLHLIG
 301 TASTWFLDDI AFYSQNLFOK DIFSALGWIP PAQTMNALLEE VYIARAQTL
 351 IALCSTVPGY WFTVAFDRI GRFAIQLMGF FEMTVFMFAL ALPHYHWTLK
 401 DNRIGFVVMY SEFTFEFANFG PNATTFVPA EIFPARLRST CHGISAAAG
 451 AGAMVGAFGF LYAAQPTDPT KTDAGYPPGH WCEELVDRPW LCNFLGMLFT
 501 FLVPESKGS LEMDSRENEG EETVAELIRA TSGRTVAV

!!AA_SEQUENCE 1.0
 P1:T01124 - probable phosphate transporter At2g32830 [Imported] - Arabidopsis thaliana
 N:Alternate names: hypothetical protein F2AL7.3
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 02-Feb-2001
 C:Accession: T01124; T00785; A84738
 R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, S.; Mason, T.M.; Kerlavage, A.R.; Adams, M.D.; Somerville, C.R.; Venter, J.C.
 submitted to the EMBL Data Library, December 1997
 A:Description: Arabidopsis thaliana chromosome II BAC T1L14 genomic sequence.
 A:Reference number: Z14209

A:Accession: T01124
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-542 <ROU>
 A:Cross-references: EMBL:AC003033; NID:g2702261; PID:g2702279
 A:Experimental source: cultivar Columbia
 R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, S.; Mason, T.M.; Keriavage, A.R.; Adams, M.D.; Somerville, C.R.; Venter, J.C.
 Submitted to the EMBL Data Library, February 1998
 A:Description: Arabidopsis thaliana chromosome II BAC F24L7 genomic sequence.
 A:Reference number: Z14204
 A:Accession: T00785
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-542 <ROW>
 A:Cross-references: EMBL:AC003974; NID:g2914688; PID:g2914691
 A:Experimental source: cultivar Columbia
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carreira, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: AB4420; MUID:20083487
 A:Accession: AB4738
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-542 <STO>
 A:Cross-references: GB:AE002093; NID:g2702279; PID:AA891982.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: AT2g32830
 A:Map position: 2
 A:Introns: 102/1; 259/3
 A:Note: T21L14.23; F24L7.3
 C:Superfamily: probable inorganic phosphate transport protein PH084

 T01124 Length: 542 May 30, 2002 09:26 Type: P Check: 992 ..
 1 MAKKGKVLN ALDAKTONY HFTAVIAGM GFTDAYDLF SISLVTKLG
 51 RIYHVDSK KPGLPPNVA AAVNGVAFGC TLAGOLFPGW LGDKLGKRV
 101 YGTLMLMVL CSIGSLSTG HSANGVMATL CFFRWLGFG IGGDPLSAT
 151 IMSEYANKKT RGAFAVFA MGFGLAGG IVSLVSTF DHAFKAPTE
 201 VDPGVSTVPQ ADVVWRIVL FGAIPALTY YWRKMPETA RYALVARNT
 251 KQAADMSKV LQVDLIAEE AOSNSNSNP NTFGLFIRE FARRHGLHL
 301 GTTTWFLD IAVYSSNLFQ KDIYTAIGNI PAETMAIH EYTVSKAQT
 351 LIALGTVPG YWFTVAFIDI LGRFIOIMG FIFMTIFMFA LAIPYDWRH
 401 RENNIGLIM YSLTMEFANF GPNATTFVP AEIPPARLS TCHGISASG
 451 KGAIVGAFG FLVAAOSSDS EKTDAGYPPG IGVNSLML ACVNLGIVF
 501 TLIVPSKSK SLSEISREDE EOGSGDTVE MTVANSGRKV PV

 11AA_SEQUENCE 1.0
 P1:T07892 - Probable inorganic phosphate transport protein P1 - barrel medic
 C:Species: Medicago truncatula (barrel medic)
 C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 23-Mar-2001
 C:Accession: T07892
 R:Lin, H.; Trieu, A.T.; Blaylock, L.A.; Harrison, M.J.
 M1: Plant Microbe Interact. 11, 14-22, 1998
 A:Title: Cloning and characterization of two phosphate transporters from

Medicago truncatula roots: regulation in response to phosphate and to
 colonization by arbuscular mycorrhizal (AM) fungi.
 A:Reference number: Z16196; MUID:98086876
 A:Accession: T07892
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-537 <LIU>
 A:Cross-references: EMBL:AF000354; NID:g2465526; PID:AA81346.1; PID:g2465527
 C:Genetics:
 A:Gene: P1
 C:Superfamily: probable inorganic phosphate transport protein PH084

 T07892 Length: 537 May 30, 2002 09:26 Type: P Check: 33 ..
 1 MSGELGVINA LDVAKTQLYH FTTIYIAGMG FETDAYDLFC ISLVTKLLGR
 51 IYTEPNPTR PGILPESAQS AVTGVAVGT LAGOLFPGWL GDKIGRRKV
 101 GLTILMVC SVSGSLFSGS SPKSVATLC FFRWLGFGI GGDYPLSATI
 151 MSEYANKKTR GAFIAVFAV OGRGILGGGI VALIVASIFD HKYKVPFEE
 201 NPATSLVPQ FDYVWRILIM FGAIPALTY YWRKMPETA RYALVAKNA
 251 KQAADMSKV LQVELEVEEE KYQKMTSDKR NSYGLFSKOF AARHGLALFG
 301 TCSWFLDID AFYSQNLFOK DIFSAIGWIP PAKENAIHE VYKIRAOVL
 351 IALCSTVPGY WFTVAFIDHM GRPAIDMGF FFMVFMFGL AIPYDHSKE
 401 ENRIGFVMY SLTFEFSNFG PNAATFVPA EIFPARLST CHGISAAAGK
 451 AGAIVGARGF LVAAQSKDPT KTDKGYPTGI GIKNSLIMLG VINVGMLCT
 501 LIVPSKSKS LDELSCNEG EGAETEDRGS PRFENVA

 11AA_SEQUENCE 1.0
 P1:T07894 - Probable inorganic phosphate transport protein P2 - barrel medic
 C:Species: Medicago truncatula (barrel medic)
 C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 23-Mar-2001
 C:Accession: T07894
 R:Lin, H.; Trieu, A.T.; Blaylock, L.A.; Harrison, M.J.
 M1: Plant Microbe Interact. 11, 14-22, 1998
 A:Title: Cloning and characterization of two phosphate transporters from
 Medicago truncatula roots: regulation in response to phosphate and to
 colonization by arbuscular mycorrhizal (AM) fungi.
 A:Reference number: Z16196; MUID:98086876
 A:Accession: T07894
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-533 <LIU>
 A:Cross-references: EMBL:AF000355; NID:g2465528; PID:AA81347.1; PID:g2465529
 C:Genetics:
 A:Gene: P2
 C:Superfamily: probable inorganic phosphate transport protein PH084

 T07894 Length: 533 May 30, 2002 09:26 Type: P Check: 2173 ..
 1 MSGELGVINA LDVAKTQLYH FTTIYIAGMG FETDAYDLFC ISLVTKLLGR
 51 IYTEPNPTR PGILPESAQS AVTGVAVGT LAGOLFPGWL GDKIGRRKV
 101 GLTILMVC SVSGSLFSGS SPKSVATLC FFRWLGFGI GGDYPLSATI
 151 MSEYANKKTR GAFIAVFAV OGRGILGGGI VALIVASIFD HKYKVPFEE
 201 NPATSLVPQ FDYVWRILIM FGAIPALTY YWRKMPETA RYALVAKNA
 251 KQAADMSKV LQVELEVEEE KYQKMTSDKR NSYGLFSKOF AARHGLALFG
 301 TCSWFLDID AFYSQNLFOK DIFSAIGWIP PAKENAIHE VYKIRAOVL
 351 IALCSTVPGY WFTVAFIDHM GRPAIDMGF FFMVFMFGL AIPYDHSKE

401 ENRIGEVYIT SLTEFFANFG PNATTFVVA EIFPARLRST CHGISAAAG
 451 AGAIVGAFG LYAAOSKDPY KTDKGYPTGI GIKNSLIMLG VINFGMLCT
 501 LVPESNGKS LEDLSRENEG EGAEATEQEG SRV

!!AA_SEQUENCE 1.0
 P1:T05714 - probable inorganic phosphate transport protein - tomato
 C:Species: Lycopersicon esculentum (tomato)
 C>Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 08-Oct-1999
 C:Accession: T05714
 R:Jiu, C.; Muchhal, U.S.; Mukatira, U.; Kononowicz, A.K.; Raghochama, K.G.
 Submitted to the EMBL Data Library, September 1997
 A:Description: Tomato phosphate transporter genes are differentially regulated
 in plant tissues by phosphorus.
 A:Reference number: Z15426
 A:Accession: T05714
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-538 <LIU>
 A:Cross-references: EMBL:AF022873; NID:g2570522; PIDN:AB82146.1; PID:g2570523
 C:Genetics: Pn1
 A:Gene: Pn1
 C:Superfamily: probable inorganic phosphate transport protein PH084

T05714 Length: 538 May 30, 2002 09:26 Type: P Check: 4793 ..

1 MANDLOVINA LDVAKTQLYH FTAIVAGMG FFTDAYDLFC ISWVKILGR
 51 LYHHHGALK PGSLEPPNVA AVNGVAFGCT LAGOLFEGWL GDMGKKRYV
 101 GMTLMIVIC SIASGLSFQH TPKGWMTLTC FFRFWLGFEGI GGDYPLSATI
 151 MSEYANKKTR GAFIAVAFAM OGFGILAGM VAIIVSAFK GAPPAPAYEV
 201 DAGSTYPOA DEFWRIILMF GAIRPAGLTYV WRMKMPETAR YTLVAKNKL
 251 QAAANDSKVL QVEIEAPEK VTAISEAKGA NDGGLFTKEF LRRHGLHLG
 301 TASTWFLIDI AFYSONLFQK DIFSAIGWIP PAQTMNALLEE VYKIRAOQL
 351 IALCSTYPGY WFTVAFIDKI GRFAIQLMGF FEMTVEMFAL AIYHHMTLK
 401 DHRIGEVYMT SFTFFPANFG PNATTFVVA EIFPARLRST CHGISAAAG
 451 AGAMVGAFFG LYAAOPTDPT KTDAGYPPGI GVRNSLIIVG CVNPLGMLFT
 501 FLVPESNGKS LEDLSRENEG EETVAEIRA TSGRTVPV

!!AA_SEQUENCE 1.0
 P1:T07164 - probable inorganic phosphate transport protein - tomato
 C:Species: Lycopersicon esculentum (tomato)
 C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 20-Jun-2000
 C:Accession: T07164
 R:Tabiri-Alaoui, A.; Avrova, A.; Antoniw, J.F.
 Submitted to the EMBL Data Library, January 1998
 A:Description: A phosphate transporter from tomato roots up-regulated during
 arbuscular mycorrhiza colonization.
 A:Reference number: Z15971
 A:Accession: T07164
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-538 <TAH>
 A:Cross-references: EMBL:Y16125; PIDN:CAA76075.1
 A:Experimental source: cultivar Early Mech; root
 C:Superfamily: probable inorganic phosphate transport protein PH084

T07164 Length: 538 May 30, 2002 09:26 Type: P Check: 5788 ..

1 MANDLOVINA LYVAKTQLYH FTAIVAGMG FFTDAYDLFC ISWVKILGR
 51 LYHHHGALK PGSLEPPNVA AVNGVAFGCT LAGOLFEGWL GDMGKKRYV

101 GMTLMIVIC SIASGLSFQH TPKGWMTLTC FFRFWLGFEGI GGDYPLSATI
 151 MSEYANKKTR GAFIAVAFAM OGFGILAGM VAIIVSAFK GAPPAPAYEV
 201 DAGSTYPOA DEFWRIILMF GAIRPAGLTYV WRMKMPETAR YTLVAKNKL
 251 QAAANDSKVL QVEIEAPEK VTAILEKOGA NDGGLFTKEF LRRHGLHLG
 301 TASTWFLIDI AFYSONLFQK DIFSAIGWIP PAQTMNALLEE VYKIRAOQL
 351 IALCSTYPGY WFTVAFIDKI GRFAIQLMGF FEMTVEMFAL AIYHHMTLK
 401 DHRIGEVYMT SFTFFPANFG PNATTFVVA EIFPARLRST CHGISAAAG
 451 AGAMVGAFFG LYAAOPTDPT KTDAGYPPGI GVRNSLIIVG CVNPLGMLFT
 501 FLVPESNGKS LEDLSRENEG EETVAEIRA TSGRTVPV

!!AA_SEQUENCE 1.0
 P1:T05724 - probable inorganic phosphate transport protein PT2 - tomato
 C:Species: Lycopersicon esculentum (tomato)
 C>Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 08-Oct-1999
 C:Accession: T05724
 R:Jiu, C.; Muchhal, U.S.; Mukatira, U.; Kononowicz, A.K.; Raghochama, K.G.
 Submitted to the EMBL Data Library, September 1997
 A:Description: Tomato phosphate transporter genes are differentially regulated
 in plant tissues by phosphorus.
 A:Reference number: Z15426
 A:Accession: T05724
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-528 <LIU>
 A:Cross-references: EMBL:AF022874; NID:g2570524; PIDN:AB82147.1; PID:g2570525
 C:Genetics: PT2
 A:Gene: PT2
 C:Superfamily: probable inorganic phosphate transport protein PH084

T05724 Length: 528 May 30, 2002 09:26 Type: P Check: 3242 ..

1 MAVGDNDNNN LOVINALDLA KTQLYHPTAI VIAQMGFTD AYDLFSISLV
 51 TKLLGRLYYT KPDLLKPGTL PPAVSASVYG VALVGTLAGQ LFFGWLGDKM
 101 GRKKVYGMTL VLMVCSVAS GLSPGSTPKG VMTLLCFEFP WLFGIGGDY
 151 PLSATINSEY ANKKTGAPL AAVFAMQFG ILFGIYALI TANGPDHAYR
 201 SPTEENAL STYPOSDIYM RIILMFGSLP AALTYWRMK MPETARYTAL
 251 VANDAKRAAQ DMGKVLQVEI ESEAKIBOI SRNETNOFGL FSWEPVRRHG
 301 LHLFGTSTW FLIDIAFYSQ NLFQKDVSA VGMIPKAPTM NAOVEYYKTA
 351 RAOTLLALCS TYPGYWTYVA FIDITGRPAI QLMGFFMTV EMEALIPPH
 401 HMTLEANRIG FIVMSLTFP FANFGPNATF FVVAEILEPA RLSTCHGIS
 451 AAGAKAGAIY GAVGFYLAQ SKDPKKTGAG YPACIGIKNS LIYLGGINL
 501 GMLCTFCVPE PKGSLSEAS QETITGEA

!!AA_SEQUENCE 1.0
 P1:T07808 - probable inorganic phosphate transport protein - Madagascar
 periwinkle
 C:Species: Catharanthus roseus (Madagascar periwinkle)
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Jun-2000
 C:Accession: T07808
 R:Kai, M.; Masuda, Y.; Kikuchi, Y.; Osaki, M.; Tadanu, T.
 Soil Sci. Plant Nutr. 43, 227-235, 1997
 A>Title: Isolation and characterization of a cDNA from Catharanthus roseus
 which is highly homologous with phosphate transporter.
 A:Reference number: Z07788

A:Accession: T07808
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-542 <RAI>
 A:Cross-references: EMBL:AB004809
 C:Genetics:
 A:Gene: P111
 C:Superfamily: probable inorganic phosphate transport protein PH084
 T07808 Length: 542 May 30, 2002 09:26 Type: P Check: 2442 ..

1 MAKQLOVLN ALDVAKTQW HTTAIVICM GFETDAYLF CISTVTKLG
 51 RIYHVDCAV KPSTLPNNVS AAVNGVAFCG TLAGOLFPGW LGDKMKRRV
 101 YGWTLMMLVI ASIASGSGFG DKPKAVMNL CFFRWLFGF IGGDYPISAT
 151 IMSEYANKKT RGAFLAVFA MGFGLITGG IFALIVSAF EAKRPATYQ
 201 DGAAVSTVPE ADYVWRITLM FGAIPALTY YWRMKMPETA RYTALVAKNA
 251 KQANMSKV LOYLEAEQD KYEKFAQERA NTFGLTFTEF LKRHGLHLG
 301 TATTWFLDI AFYSQNLFOK DIFSAIGWIP PAQTMAIEE VFRIRAOVL
 351 IALCSTVPGY WFTVFLIDRI GRFIIQMMGF FFWTVFMPAL AIPYNHTHK
 401 DNRGFIWY SLTFFFNFG PNATTFVYPA EIFPARLST CHGSAAGK
 451 AGAIIAGRF LYAAPSDPS KTDKGYPCI GKNALIVLG CVNLFQVFT
 501 FLVPEAKGS LEEVSKENE EVENGTELRO QSGHDTFTVP VL
 IIAA_SEQUENCE 1.0
 P1:T47629 - phosphate transport protein - Arabidopsis thaliana
 N:Alternate names: protein T5N23.60
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 19-May-2000
 C:Accession: T47629
 R:Obermaier, B.; Othenwaeler, B.; Duchemin, D.; Zeltner, K.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat M.; Mewes, H.W.; Lemcke, K.;
 submitted to the Protein Sequence Database, March 2000
 A:Reference number: 224463
 A:Accession: T47629
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-535 <OBE>
 A:Cross-references: EMBL:AL138650
 A:Experimental source: cultivar Columbia; BAC clone T5N23
 C:Genetics:
 A:Map position: 3
 A:Note: T5N23.60
 C:Superfamily: probable inorganic phosphate transport protein PH084
 T47629 Length: 535 May 30, 2002 09:26 Type: P Check: 3263 ..

1 MAGQOLVLN ALDVAKTQW HTTAIVICM GFETDAYLF CISTVTKLG
 51 RIYHVDCSE KPSTLPNNVS AAVNGVAFCG TLAGOLFPGW LGDKMKRRV
 101 YGWTLMWVL CSIASGSGFG SNPKTYMTL CFFRWLFGF IGGDYPISAT
 151 IMSEYANKKT RGAFLAVFA MGFGLITGG IFALIVSAF EAKRPATYQ
 201 IDALASTVPO ADYVWRITLM VGALPAAMTY YSRKMPETA RYTALVAKDA
 251 KLASNSKV LOYLEAEQD GTEKSNSTFG LPSKEFMKRH GLHLGTST
 301 WFLDIAFYS QNLFQKDFIS AIGWIPPAOT MNAIOEVKRI ARQTLTALC
 351 STVGCIWFTV AFIDVIGRFA IOMMGFEWTF VFMFLAIPY DHWTNENRI

401 GFVAMYSLTF EFANFGPNAT TEVPAEIRP ARFRSTCHGI SAAAGKIGAM
 451 VGAFGEFLYA QSPDKTEH GYPPGIVKN SLIVGVNL LGWFTLLVP
 501 ESKGSLEEM SGENQNDSS SSSNNNSNN AVSTA
 IIAA_SEQUENCE 1.0
 P1:T34457 - hypothetical protein T19H12.10 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T34457
 R:Davidson, S.
 submitted to the EMBL Data Library, April 1997
 A:Description: The sequence of C. elegans cosmid T19H12.
 A:Reference number: 221528
 A:Accession: T34457
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-533 <DAV>
 A:Cross-references: EMBL:U97009; PIDN:AA69033.1; GSPDB:GN00023; CESP:T19H12.10
 A:Experimental source: strain Bristol N2; clone T19H12
 C:Genetics:
 A:Gene: CESP:T19H12.10
 A:Map position: 5
 A:Introns: 50/3; 73/3; 126/1; 156/1; 475/2
 C:Superfamily: glucuronosyltransferase
 T34457 Length: 533 May 30, 2002 09:26 Type: P Check: 9544 ..

1 MRLTFSCLC FLKVNVAAR ILIFNDFGF SHKFEVSKLA DIADHGOEV
 51 TLFOFHLAL KNLDGIYKRN NIEIINYPD HYDDLKLET QTFPDPWDSQ
 101 LMNPVLMAF MLPRILGGEF KTTIOLIND KEILKLKDK KFDVAISETF
 151 ELTGWYMSHF LGVPCIPILS AVRLDIFNEA FGQSSAFGYL TQOQSKLAPD
 201 AGFLDLRLVD YRDFEFSKMAF RGMAYQONDY IEKAGHYVP YWKDLVKEAP
 251 VYMNNSNYL DEAVPTTATI VHGIGITTL EKMNVDALP EBYIILKEK
 301 ETVYLISFGS VINSYEMPN FKAGLIKVPE SLDPVIFLWK YEIDDLFEOK
 351 KLPKNVHLK WVPQPSLLAD KRVKLFVTHG GIGSTMEVAY TGRPALWPI
 401 FGDDPMNMD LARIGAIAY DKFDLYDGKK LLETVRDLVT NPKKEQAKE
 451 LLDVLTNPI DRYNNLMKHL EFAIKPNLR SQIPEINOVG PIAHYLDVI
 501 VFLFVSILT AVISFOIVCR ILSRIISKV KSD
 IIAA_SEQUENCE 1.0
 P1:S56753 - interferon regulatory factor 3 - chicken
 C:Species: Gallus gallus (chicken)
 C>Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
 C:Accession: S56753
 R:Grant, C.E.; Vasa, M.2.; Deeley, R.G.
 Nucleic Acids Res. 23, 2137-2146, 1995
 A:Title: GIRF-3, a new member of the interferon regulatory factor (IRF) family that is rapidly and transiently induced by dsRNA.
 A:Reference number: S56753; MUID:95334365
 A:Accession: S56753
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-491 <GRN>
 A:Cross-references: EMBL:U20338; NID:9790580; PIDN:AAA86995.1; PID:9790581
 C:Superfamily: lymphoid-specific interferon regulatory factor
 S56753 Length: 491 May 30, 2002 09:26 Type: P Check: 3644 ..

1 MAALDSEDA OKLRFPMLL NAVSSGLYRG LCWIDPDRRI FRIRWKINAR
 51 KDVTSSDVEI FKAMAKASGR YEGNAEDPAK WKTNRCALR STMFMLEED

101 RSKNDPPHK VYAVASGVPN DRGSGCPYAG ALQOQPOLLL NHHDLLAENT
 151 PUDSTEGVAA AALTOYDIDL LQSVLQHCNI SALGSOPTLW AHNGDALPED
 201 ALLPGQDGC LPPGFQFQWR QLEEPILLGN QPLTGGGCGQ DGAGALPVSE
 251 ECALPASPRA EELLPOSANP APPRAGDIG GLPPLDITTI YYGKNWYOZ
 301 QVDSRCVLA YOPUDPAVAE QRLVLEPSPA SLDPDRQRRY TEDLLEVAGL
 351 RLEQFAGQLL ATRLKCKRYE WALSOQLBEG EPPNLHLRD QETTFIDERY
 401 FCELRNFRD SRRESPPDT IFLCGQCFPS STRKESKLI LKLVYQFCE
 451 YWEQVORG ASSLSGNVS LQSDSFNLF ELIBQYHMOT D

!!AA_SEQUENCE 1.0
 PI:S73432 - MG096 homolog D09.orf518 - Mycoplasma pneumoniae (strain ATCC 29342)
 C:Species: Mycoplasma pneumoniae
 A:Variety: ATCC 29342
 C:Date: 26-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
 C:Accession: S73432
 R:Himmelfreid, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
 Nucleic Acids Res. 24, 4420-4449, 1996
 A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.
 A:Reference number: S73327; MUID:97105885
 A:Accession: S73432
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-518 <HIM>
 A:Cross-references: EMBL:AE000013; GB:U00089; NID:91673762; PIDN:AAB95754.1; PID:91673763
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
 C:Genetics:
 A:Genetic code: SGC3
 C:Superfamily: hypothetical protein MG096

S73432 Length: 518 May 30, 2002 09:26 Type: P Check: 4533 ..

1 MNELOTKVNE ANNIPVEAF KVPKPEKLF GFVNQGFEPK LMPKGLINID
 51 NVASLEQYS LKQASLKDPD ILLEKNDIV LEHKVRNFA LOHFETTVY
 101 GTEGELNLOP ALQASTTNFS SLEELQASFS KTGDNLTAQL FPKPTVKLV
 151 SGENDLTTHIA QTAIGSELEF SRVDLSASII NSEATLKTAE ATFTQVLANP
 201 FKAEREKALA IKKAEERIK KELEEQKKRO EELSKOORDK EALQSLMANF
 251 QEFISTWQO GNDVKQKQEOF IQALEAFAST NMNEVENLLI AGFRSAIQTY
 301 YKDKADQSO NAKIAFGEKG IQFPKSGPGL DGIFMSDFLR GNULTGNARHD
 351 LKTKKEVAN TQCKDAQGND KKASINNOAK QNNFPPROVN PNDPSEVEL
 401 KTEGSGILYP GARPLNLFQS LGIPNDWKE MSVKFVLGDK TPQWITADKPD
 451 YEGSLFKFER NQLKFTPHVK EHVHVENKQF MEKLKSONLH NIELATGATK
 501 PPVVDLASYL HYLILNHK

!!AA_SEQUENCE 1.0
 PI:S73923 - MG288 homolog G07_orf417 - Mycoplasma pneumoniae (strain ATCC 29342)
 C:Species: Mycoplasma pneumoniae
 A:Variety: ATCC 29342
 C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
 C:Accession: S73923
 R:Himmelfreid, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
 Nucleic Acids Res. 24, 4420-4449, 1996
 A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma

pneumoniae.
 A:Reference number: S73327; MUID:97105885
 A:Accession: S73923
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-417 <HIM>
 A:Cross-references: EMBL:AE000058; GB:U00089; NID:91674291; PIDN:AAB96245.1; PID:91674299
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
 C:Genetics:
 A:Genetic code: SGC3
 C:Superfamily: hypothetical protein MG096

S73923 Length: 417 May 30, 2002 09:26 Type: P Check: 7095 ..

1 MGFHAYQNN DFKANDPAL NISTINPAIL QELONSPDQ GSPLTAGLFY
 51 KYSVKNLTSG TNDLTITANT ALGENIQRQ VLTQSIIRP RLEPAATOYK
 101 QDILAPPAKE RQALALQHLK EIEEAKORAE QLKEQOEAE KRRQEEVKVY
 151 AETQGFNDL TSAQKFKRYW LKQGDVTKK VELIQALKSS FFRNOKRTFN
 201 FLNAGFTAI DMYNOEKNN TTAKNNAFGK NGIQFPVAGF QGTYMSQWLR
 251 DELSGKTDIK LNKSLSYON ENKNSINWN KQKRIEIKOV KPNYSFEIN
 301 LKTYGSYNS LMYLIGAAG GIPTSMSGM DMKRTIVGDL DSGIVTKQDY
 351 PGSEFETED KLMFTLHVQO QIKKEQGFN NLKGQSLDN LDLRGTTPK
 401 PVDVLAAYL FVILTPAK

!!AA_SEQUENCE 1.0
 PI:B72719 - probable bacterioferritin comigratory protein APE0291 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: B72719
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takahashi, M.; Sekine, M.; Baba, S.; Ankel, A.; Kosugi, H.; Hosoyama, A.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, M.; Masuda, S.; Aoki, K.; Kubota, K.; Nakamura, Y.; Nomura, N.; Sako, Y.; Kikuchi, H.
 DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix K1.
 A:Reference number: A72450; MUID:99310339
 A:Accession: B72719
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-110 <KAW>
 A:Cross-references: DDBJ:AP000059; NID:95103911; PIDN:BAA79246.1; PID:95103930
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE0291
 C:Superfamily: bacterioferritin comigratory protein; alkyl hydroperoxidase c22 protein homology

B72719 Length: 110 May 30, 2002 09:26 Type: P Check: 4613 ..

1 MUSVGDPAID IEIQLIDGST IRLSQLRGSR VLYFYPKAP TPQCTREATIG
 51 ENGLYEEFK LGAEVIGVSM DPGGNRRFA QNTGVRRFLA SVEGGRAPFS
 101 FMGAQRPTD

!!AA_SEQUENCE 1.0
 PI:G70077 - hypothetical protein yxiX - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: G70077

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessières, P.; Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans, A.; Braun, M.; Brignelli, S.C.; Bron, S.; Broutier, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani, J.J.; Conerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denizot, F.; Devine, K.M.; Dueterhoeft, A.; Ehlich, S.D.; Emerson, P.T.; Entlan, K.D.; Errington, J.; Fader, C.; Ferrari, E.
 Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gallen, N.; Ghm, S.Y.; Glaeser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.; Guiseppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Hatwood, C.R.; Henault, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.; Kasahara, Y.; Kjaer-Blanchard, M.; Klein, C.; Kobayashi, Y.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.
 A:Authors: lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno, M.; Mostl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Potwolk, S.; Prescott, A.M.; Presecan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.; Ray, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, E.
 A:Authors: Schleich, S.; Schnoeller, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, S.J.; Serro, P.; Shiu, B.S.; Soldo, B.; Sorokin, A.; Taccini, E.; Takagi, T.; Takahashi, H.; Takemaru, K.; Takeuchi, M.; Takashi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vandenberg, M.; Vannier, F.; Vassartelli, A.; Varti, A.; Wambuit, R.; Wedler, E.; Wedler, H.; Welterneger, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; MUID:98044033
 A:Accession: G70077

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-153 <KUN>
 A:Cross-references: GB:299124; GB:AL009126; NID:92636442; PIDN:CAB15950.1; PID:92636460
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yxk
 C:Superfamily: *Bacillus subtilis* hypothetical protein yxk
 G70077 Length: 153 May 30, 2002 09:26 Type: P Check: 7683 ..

1 MEITISISIG NLDMIELKPD QYMACLEL AESFYFMAG LAYDRIMQV
 51 ITTGSFIEDL SEYFEGHAYK VTKLANREFH FOSILOEADR DIADFLFLIA
 101 SINDVFLIT DPDPKSYFS EGKIQCLTDS GERIWFED AVDIYMGGE
 151 SYK

11AA_SEQUENCE 1.0
 PI:T29895 - hypothetical protein F38A5.13 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Apr-2000
 C:Accession: T29895

R:Fulton, B.; Stellyes, L.
 submitted to the EMBL Data Library, September 1996
 A:Description: The sequence of C. elegans cosmid F38A5.
 A:Reference number: Z20705
 A:Accession: T29895
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-589 <FULT>
 A:Cross-references: EMBL:U70854; PIDN:AAB09157.1; GSPDB:GN00022; CESP:F38A5.13
 A:Experimental source: strain B1stcol N2; clone F38A5
 C:Genetics:
 A:Gene: CESP:F38A5.13
 A:Map position: 4
 A:Introns: 28/3; 121/1; 299/3; 345/3; 466/3

C:Superfamily: mouse Id-associated protein 1; dnaJ amino-terminal homology; myb DNA-binding repeat homology

T29895 Length: 589 May 30, 2002 09:26 Type: P Check: 5664 ..

1 MTGNIQVAI YGFARSRAF EPAGICYETR LIRDKLTLCH CTPLPKKNSP
 51 VAKPYTEQOR KKSSEPAEL RKDLFDADNE KKEKYLKMD PNDCKNODHY
 101 KYLGSLKLM QATSEIRFC YRQVLKLNH DKKHGIGYA EKEEFTCTT
 151 KAYEYQMSD VKROAFDSVD HKENDIIPNE KSIHNHFN ELAPYQUNLS
 201 RMSNIKPYE LKSPDATRED VENEYDFWN FQSWREFSYL DEEDKERGED
 251 RYREREMEQ NKAERERRR EAKRIKLY DIAVADPPI IKFKEQQAK
 301 KDKAKEDQOR AIREKQEAID REKREKEAE AKQEKADRK AKEREREKK
 351 ERDIKKAKMS QQRRLKLLA DEAGHWTEPN RDKLTEMERI ERICIGFTVD
 401 QARELCEKYE SLSTASEIQT ALTDALILK EAGAATVTT EDKNNENKQ
 451 ADKETWSEE IQLVKASNT PPQETERNV QIADYNEHR KDSTGLPPXT
 501 EKQVIOCKA VQTMNVKLPS TTQNLCTAL PDEDVMSATE OKTLDEAIK
 551 HKSDPERWE KISTEVGRKS KKACIRRFKY LVQMYANKK

11AA_SEQUENCE 1.0
 PI:T19595 - hypothetical protein C31A11.5 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
 C:Accession: T19595

R:McMurray, A.
 submitted to the EMBL Data Library, December 1996
 A:Reference number: Z19149
 A:Accession: T19595
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-705 <WILL>
 A:Cross-references: EMBL:283218; PIDN:CAB05687.1; GSPDB:GN00023; CESP:C31A11.5
 A:Experimental source: clone C31A11
 C:Genetics:
 A:Gene: CESP:C31A11.5
 A:Map position: 5
 A:Introns: 23/3; 86/1; 133/3; 287/1; 381/1; 474/2; 547/2; 579/1; 627/3; 656/2
 C:Superfamily: *Caenorhabditis elegans* hypothetical protein F09B9.1
 T19595 Length: 705 May 30, 2002 09:26 Type: P Check: 676 ..

1 MKIAPLFFFI VCLFATVCE NMKEVFLAKA AOKKITGVSE QCANTEWQ
 51 KSLKNVAELS AECLIEOKCT KEELKTIEDN FVAVQYDAM GKIPPLGLFQ
 101 LPILFDGSYQ ECEIRSGKKY ATNYCYWVLM PGKNATCHMS DGLPTTFEER
 151 GAVCMYSSS EODLEPTVNO VSDQPTACA AECSSFPVAK TPAPGFISSF
 201 MAVMIGIALI ATVIDYLKDA LKREDEKRED SRLLQILTF SLMTNAELL
 251 SVKEQKPGFI KCLDCIRFLS MLWVVTGHTF SYLTPTDQIE SILPFGRFM
 301 NNLVNMAFYS VDTFFLSGL VVSLTFPKTK LKYSQKSPK TWILFYVIRY
 351 LRLTPPLMFE LGFEVYVGYK FQGPVASQL NQONGEVDC QYWNKNLTY
 401 INNLMSGDQ CYGLTWYIGA DTQYLIVAPI FLGLTFPSFA IGTALLTAAT
 451 IGSVITVYL FSTYDLPADF FGNQDATHFY DMIIYIPWR CPPIYVGLV
 501 GYLATYGRK KRLRLMALAV TGMIVAFSLG ALCTFSTYDY DNKVMKSFIS

551 RAYYNSRL AMSFALSWI VANHMGCP IDAFKSHPM QPGRISTYCA
601 YVHYVLYM YLMIGDASIH FYSSFOIEM YAVPTVLSY IFAFWSCLP
651 EIDPLKLEKM LIELIGCAR DNRREDIEKO KTLKTKENE LMAVEETVOS
701 TNEKF

!!AA_SEQUENCE 1.0
P1:B81828 - hypothetical integral membrane protein NMA1989 [imported] -
C:Species: Neisseria meningitidis (strain Z2491 serogroup A)
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: B81828
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morelli, G.; Basham, D.; Brown, D.; Chillingworth, T.; Davies, R.M.; Davis, P.; Devlin, K.; Feltwell, T.; Hamlin, N.; Holtrold, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, M.A.; Rutherford, K.M.; Simmonds, M.; Skelton, J.; Whitehead, S.; Spratt, B.G.; Barrell, B.G.
Nature 404, 502-506, 2000
A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556
A:Accession: B81828
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-172 <PAR>
A:Cross-references: GB:AL162757; GB:AL157959; NID:97380371; PIDN:CAB85209.1; PID:97380620; GSPDB:GN00124; NMA5P:NMA1989
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1989
C:Superfamily: Neisseria meningitidis hypothetical protein NMB1733

B81828 Length: 172 May 30, 2002 09:26 Type: P Check: 492 ..

1 MKDSIMDAFD FFMGNFKKI ATEGLDRKY FIFQIYCNL SAISSCLMK
51 IIRYITLTYI LIAVIGING FLNRKVLST VKNKKFLSD MPLEMPFFF
101 VEKKIPIILAR FLNRGMKWL LPSFCLVAI CFSIYILNEE GGMMPYFACY
151 WFGICYPVST YVFLSRDYK HI

!!AA_SEQUENCE 1.0
P1:D81049 - hypothetical protein NMB1733 [imported] - Neisseria meningitidis (strain MC58 serogroup B)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: D81049
R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Ketchum, K.A.; Hood, D.W.; Peden, J.F.; Dodson, R.J.; Nelson, W.C.; Gwin, M.L.; Debey, R.; Peterson, J.D.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Mason, T.; Ciecko, A.; Parksey, D.S.; Blair, E.; Ciltone, H.; Clark, E.B.; Cotton, M.D.; Uffereback, T.R.; Khouri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignant, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappapoli, R.; Venter, J.C.
A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: D81049
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-167 <TEET>
A:Cross-references: GB:AE002523; GB:AE002098; NID:97226960; PIDN:AAF42078.1; PID:97226984; GSPDB:GN00119; TIGR:NMB1733
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1733
C:Superfamily: Neisseria meningitidis hypothetical protein NMB1733

D81049 Length: 167 May 30, 2002 09:26 Type: P Check: 3936 ..

1 MDARDFEMOG NFKKIATEOG LDRKYFIFQY IVCNLSAIS FCLMKIIRHTY
51 TILYITLAVI GIINGFLRS KVLSTYKNO KFLSDMFLE MPFFVEYKI
101 PIARFLNRG MKWLYLPSFF CLVAICFSIY ILNEBGWMP YFACWYEGLC
151 YPVSTYVFL SRDYKHI

!!AA_SEQUENCE 1.0
P1:A81923 - probable membrane protein NMA0784 [imported] - Neisseria meningitidis (strain Z2491 serogroup A)
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: A81923
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morelli, G.; Basham, D.; Brown, D.; Chillingworth, T.; Davies, R.M.; Davis, P.; Devlin, K.; Feltwell, T.; Hamlin, N.; Holtrold, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, M.A.; Rutherford, K.M.; Simmonds, M.; Skelton, J.; Whitehead, S.; Spratt, B.G.; Barrell, B.G.
Nature 404, 502-506, 2000
A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556
A:Accession: A81923
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-173 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:97379424; PIDN:CAB84067.1; PID:97379505; GSPDB:GN00124; NMA5P:NMA0784
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0784
C:Superfamily: Neisseria meningitidis hypothetical protein NMB1733

A81923 Length: 173 May 30, 2002 09:26 Type: P Check: 1872 ..

1 MKDSIMDFFD FFAQGRYKKI ATEGLDRKY FIFQIYCNL SAISSCLMK
51 VIHSYITLCL LIMIGIAFG FLNRNVLST VKNKKFLSD MPLEMPFFF
101 RVYKIPILAR FLNRGMKWL LQNFCLVMI CFSISILINE ERGMMPYFIC
151 YVFGIFYPAN TYVFLSRDY KRI

!!AA_SEQUENCE 1.0
P1:S28025 - light harvesting complex II assembly factor pucc [imported] - Rhodobacter sphaeroides
C:Species: Rhodobacter sphaeroides
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 28-Jul-2000
C:Accession: S28025; T50701
R:Gibson, L.C.D.; McGlynn, P.; Chaudhuri, M.; Hunter, C.N.
Mol. Microbiol. 6, 3171-3186, 1992
A>Title: A putative anaerobic cytoporphyrinogen III oxidase in Rhodobacter sphaeroides. II. Analysis of a region of the genome encoding hemp and the puc operon.
A:Reference number: S28023; MUID:93086425
A:Accession: S28025
A:Molecule type: DNA
A:Residues: 1-459 <GTB>
A:Cross-references: EMBL:X68796; NID:946438; PIDN:CAA48701.1; PID:946441
R:Choudhary, M.; Kaplan, S.
Nucleic Acids Res. 28, 862-867, 2000
A>Title: DNA sequence analysis of the photosynthesis region of Rhodobacter sphaeroides 2.4.1.
A:Reference number: Z25222; MUID:20115911
A:Accession: T50701
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-459 <CHO>
A:Cross-references: EMBL:AF195122; PIDN:AAF24245.1

A:Experimental source: strain 2.4.1
C:Genetics:
A:Gene: pucc

S28025 Length: 459 May 30, 2002 09:26 Type: P Check: 7732 ..

1 MSRIAEHLVR IGRPLRFAD AASDLPRLK LLRLSLFOVA VGMAYLLVG
51 TLNRVMIVEL KVPASVVGIM ISLPLLPAPF RALIGFSDT HVSALGRRV
101 PWIRGTAL MGFAIMPPA LIVAGGQVA EGQPFMLGV SAAALAPLVG
151 GGVIHTQTVG LALATDLAR EDQPKVGLM YVVLISMIF ASIGFGLD
201 PYDAQILIK ISGVAAVAF LNMIALMKE PRNRAFVYK EKPERGDHW
251 REFSRENAL HGLIVIGLCT LGFGMADYIL EPYGEVLWM TVAEETRLTA
301 TFAGGLVGF WLASWVLRG FDPRLMAFLG AAAGLPGEFA IMGATEMTNV
351 WVELLGTLLV GFGGLFSGH TLTATMLAP KEQVGLALGA WGAVOATAG
401 VALAGACVLR DILQAMPDLS GYGPAPYVA VFALEAGFLF LTMIVILPL
451 RSALAAARL

11AA_SEQUENCE 1.0

P1:A64690 - competence locus E - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C>Date: 09-Aug-1997 #sequence,revision 09-Aug-1997 #text,change 08-Oct-1999

C:Accession: A64690

R:Tombl, J.F.; White, O.; Karlavage, A.R.; Clayton, R.A.; Sutton, G.G.;
Fleischmann, R.D.; Ketchum, K.A.; Klein, H.P.; Gill, S.; Dougherty, B.A.;
Neilson, K.; Queckenbush, J.; Zhou, L.; Kirness, E.F.; Peterson, S.; Loftus,
B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, K.;
Fitzgerald, L.M.; Lee, N.; Adams, M.D.; Hickey, E.K.; Berg, D.E.; Gocayne,
J.D.; Uterback, T.R.; Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman,
J.M.; Fujii, C.; Bowman, C.; Matthey, L.
Nature 388, 539-547, 1997

A:Authors: Mallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.;
Fraser, C.M.; Venter, J.C.

A>Title: The complete genome sequence of the gastric pathogen Helicobacter
pylori.

A:Reference number: A64520; MUID:97394467

A:Accession: A64690

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-417 <TOM>
A:Cross-references: GB:AE000636; GB:AE000511; NID:92314517; PIDN:AMD08401.1;
PID:92314527; TIGR:HP1361

C:Genetics:

A:Start codon: GTC

A64690 Length: 417 May 30, 2002 09:26 Type: P Check: 8156 ..

1 MCGVFLSLL AINLYLEYLN YOKLDFSKPT SLQAQILQY PKTRDQTYE
51 VIKQSKNMI FTTIKRPLK NIQYRHAQF GKIKPSCFL SAKSCFQTY
101 SFSITRKQDF KSHWRHFDIS AHENALVGNL YRALFIGDSL NKDIDRANA
151 LGINHLAIS GFHGLISVS VFELSLFYT PLOKRYEPPYR NAFYDGLV
201 WVELLGYLL LDPLSPFRA FLMGGLGFLA CFFGVRLSLF KLILACCIA
251 IALLPKLFS VGFLLSVCGV WTIFLPLKHT QIEFTSSFL MRSEQALSL
301 ALVELNLII VHAFFMFSP YOLFSPILGL IFIVFFPLSL FLHVGIGSL
351 LDRLSMPLT IPTISVPSPL WLVGHLFLT ILSARFFKVY LSMNVLSAGF
401 FLYCCYQYII MPSLIYG

11AA_SEQUENCE 1.0
P1:S74540 - hypothetical protein sl1066 - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C>Date: 25-Apr-1997 #sequence,revision 25-Apr-1997 #text,change 20-Jun-2000

C:Accession: S74540
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.;
Miyajima, N.; Hirosewa, M.; Sugita, M.; Sasamoto, S.; Kimura, T.; Hosouchi,
T.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimo, S.;
Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 3, 109-136, 1996

A>Title: Sequence analysis of the genome of the unicellular cyanobacterium
Synechocystis sp. PCC6803. II. Sequence determination of the entire genome and
assignment of potential protein-coding regions.

A:Reference number: S74322; MUID:97061201

A:Accession: S74540

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-290 <KAN>
A:Cross-references: EMBL:D90899; GB:AB001339; NID:91651650; PIDN:BAI16692.1;
PID:91651764

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June
1996

S74540 Length: 290 May 30, 2002 09:26 Type: P Check: 1310 ..

1 MPKSTVNFNF NQQAIAQLTA PMRIGVFIVM LAAMLPLVY PLSLAIADN
51 LRSIVMGIL FLIFVLLIF WSHWCYQPL SIKAYGYGL GWNROGVEL
101 LRGLGLGFSF TFGLEFIQGL LGNAVLAPAG DRLMTIIMG SLTGLGYALA
151 BELPFRGILL KELEGYGNK TSLASNAIIF AVLHRLKPLG EVITLQFP
201 ALVLLGLSG ITRRRHCDRL GHSIGLHGM VVAITVAVG QLVYTERMP
251 AMVTGIDRNP LSGVGIAGL CLLMLVNOG EKPWLKRLGF

11AA_SEQUENCE 1.0

P1:S73732 - MG306 homolog A05_Orf395 - Mycoplasma pneumoniae (strain ATCC 29342)

C:Species: Mycoplasma pneumoniae

A:Variety: ATCC 29342

C>Date: 27-Feb-1997 #sequence,revision 25-Apr-1997 #text,change 07-Dec-1999

C:Accession: S73732

R:Himmelreich, R.; Hilbert, H.; Piagens, H.; Pilkl, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996

A>Title: Complete sequence analysis of the genome of the bacterium Mycoplasma
pneumoniae.

A:Reference number: S73327; MUID:97105885

A:Accession: S73732

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-395 <HIM>
A:Cross-references: EMBL:AE000039; GB:U00089; NID:91674082; PIDN:AB96054.1;
PID:91674089

A>Note: the nucleotide sequence was submitted to the EMBL Data Library,
November 1996

C:Genetics:

A:Genetic code: SGC3

S73732 Length: 395 May 30, 2002 09:26 Type: P Check: 3483 ..

1 MSTPKSASF TRKNILAFSE FIAFLVVSV LTVFELDIK TGDVTKIINT
51 INRTNPMIL LIVIGIVYL ANIIINWVY ARRCFNAW WENYLFACV
101 QEFQIVPLS LGDPPFLRY FIKGMRKOT AVLTVSTGA FWNLAOLIT
151 WSPFVLSON YALLEQNHG EVASYWSPAG MFDVVAAIL FIFAYSKRM
201 HVLIIYGVNO FRKIRKPYL TKQIYQRFI DKAEPNKLXG LEIKRLGITI
251 EKLNLILIA VGYFVFAV FAIVKKNAT NNVIDQYSTA DIFNTINAI

301 TASNFPDVS GEGATQFWTF SFLNAFKSAV GISSQVKGCV FLMRPLSVTI
351 PALFSLCEFI GMYVQVIEF KHKRPVLEPTV NLINHEWMN KKLHN
IIAA_SEQUENCE 1.0
P1:S46356 - pupr protein - Pseudomonas putida
C:Species: Pseudomonas putida
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 15-Jun-2001
C:Accession: S46356
R:Koster, M.; van Klompenburg, W.; Bitter, W.; Leong, J.; Weisbeek, P.
EMBO J. 13, 2805-2813, 1994
A:Title: Role for the outer membrane ferric siderophore receptor PupB in signal transduction across the bacterial cell envelope.
A:Reference number: S46355; MUID:94298771
A:Accession: S46356
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-324 <KOS>
A:Cross-references: EMBL:X77918; NID:g525258; PIDN:CAA54871.1; PID:g525260
C:Superfamily: Pseudomonas putida regulatory protein pupr

S46356 Length: 324 May 30, 2002 09:26 Type: P Check: 5380 ..

1 MNGGATSTP GEVAQAMHW HLEQEPKAVS AATLAACMSW RQAHPLHEHA
51 WORTQVFAOR LREMRSPGOR PLAHALRPQ OSRRALIKOL SLIMAGAGA
101 WYLDALVQ DWRADYHSRI GEORRLTLAD GTQVQLNTDS ALNVAEDQQA
151 RRLRVGEM LITRPALADS RPLWVDPEHG RLESTLAQFN VRLHGHNTQA
201 TYVQGSVALQ PALHAYPPIL LGAGEQASFN QQGLLAQAV AAVAPWSGG
251 MLYAGQCPA AFIEDIARYR RGHLCADPAL AGLRVSGTFP LENTDKIIAA
301 VAETQLQEVQ HETRYWVLK PRMA
IIAA_SEQUENCE 1.0
P1:A11701 - hypothetical protein homolog lin2155 [Imported] - Listeria innocua (strain C11p11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: A11701
R:Glaser, P.; Fraigneul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Darvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Eutlian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maitouram, A.; Mela Vicente, J.; Ng, E.; Nordstiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund, J.; Cossart, P.
A:Title: Comparative genomics of Listeria species
A:Reference number: ABI077; MUID:21537279; PMID:11679669
A:Accession: A11701
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <GLA>
A:Cross-references: GB:A1592022; PIDN:CAAC97385.1; PID:g16414669; GSPDB:GN00178
A:Experimental source: strain C11p11262
C:Genetics:
A:Gene: lin2155

A11701 Length: 391 May 30, 2002 09:26 Type: P Check: 3736 ..

1 MKATGIVVER NPFNGHQH LNKRELTKS DVIIVWGS FGQGEPAII
51 PKWERTKML AAGVDMVIEL PVSEFATQHA IFAEESIRLL DALHDTLTF
101 GSEHGVSEDF STAAKTIVEN EAANETIQL ALGDKTISYA RAYDTAFTHL

151 EGKELDVTK PNNILGEHYA LAIQKQNPST ALQTMPREHS GYHDAEASHD
201 YIASATAIRK LLAGNLIEEA SRYLPDSSIE VLNNYRQPF LSEDFWPLLK
251 FRLIQASSDE LGICRVSESG IONRMQLAAK KAHFSDFPIE IMKTRKYSNA
301 RLQRTALQIL LNAQNTPPAE PYIRVLGMSK TGOKYISLHK KNISLPVTF
351 VSKAEPSSLK EDLRATDIYT LINGLEDYQA GDFHTPPILT L
IIAA_SEQUENCE 1.0
P1:S28472 - rfbG protein - Vibrio cholerae
C:Species: Vibrio cholerae
C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 08-Oct-1999
C:Accession: S28472
R:Manning, P.A.
submitted to the EMBL Data Library, May 1991
A:Reference number: S28467
A:Accession: S28472
A:Molecule type: DNA
A:Residues: 1-460 <MAN>
A:Cross-references: EMBL:X59554; NID:g48381; PIDN:CAA42138.1; PID:g48387
A:Experimental source: strain 017
C:Genetics:
A:Gene: rfbG

S28472 Length: 460 May 30, 2002 09:26 Type: P Check: 8871 ..

1 MTEKKRLVWV VPLPNNSSWR GEGIAQTLEN IYRNISPERK IELVSSKHA
51 EMLVGLKSN PNISVLTGCF ROKSTKKTIG YVSLNEVEKD SLMDVIANKL
101 PIIPAIFRKY GMYVQLEYL LSLYIYSHQ RGRFSSNNC RVWLPPIIP
151 YTHLGGSEKF VSFMDPEVFE YNKEPPLAE YFVAKLSKH FSNASALITQS
201 RANKVILETV MGIESKINV IYNGSPDYSE FKQOQSNLSF SEVWSKEFS
251 GASKKAFFEA LVNHQINFSEV LMRLLTKNKV SNRKIVLIST QNRPYKGFQ
301 LFLVLNELCL RRDNDYFIFT CNVPTKLKER YPSLYEINHE VYRVQYTLAA
351 SLYMSDIYV HPSNVEGGIG AYPQYASV GKPSLINTGR HYNEMAESEF
401 DVDLLSNFV NTKETVDKIE KLINSEYMR QNIDAINRLK ISWKESASNY
451 ENVEFGNENA
IIAA_SEQUENCE 1.0
P1:D82346 - rfbG protein VC0245 [Imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82346
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill, S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Baas, S.; Qin, H.; Dragol, I.; Sellers, P.; McDonald, L.; Utterback, T.; Fleischmann, R.D.; Nierman, W.C.; White, O.; Salzberg, S.L.; Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 405, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: D82346
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-463 <HEI>
A:Cross-references: GB:AE004113; GB:AE003852; NID:g9654648; PIDN:AAF93421.1; GSPDB:GN00126; TIGR:VC0245
A:Experimental source: serogroup O1, strain N16961; biotype El Tor
C:Genetics:

A:Gene: VC0245
A:Map position: 1
D82346 Length: 463 May 30, 2002 09:26 Type: P Check: 1939 ..

1 MNVTEKKKL VVVVPLPPNS SWRGEGIAQT IENIVRNISP ERKIEIVSS
51 KHAEMLVGLE KSNPNISVLT LQFRCKSTPK TIGVYSLNEV EKDSLMDLVI
101 AKLPIIRIAIF RRYGMVVSOL EYLLSLYIYS HLQRRGRSS NNCRWMLPTP
151 IIPYTHLLGC EKFSFWDPE VEEYKKEPPL TAEYFVKLLS KHPNSASAIT
201 TQSHANKDYL ETVMGIESSK INVYNGSPD YSEFKKOOSN LSFSEVMSKS
251 EFSASAKKA FEALVNHQLN FSVLWRLTLK NKYSNRKIVL ISQNRRIYKG
301 PDQLFVLIN LCLRDNDYDF IFTCNVPTKL KERRPSLYER IHEVTRVDNY
351 LHASLYLMSD IYVLPNSNVEG GLGAYPOYEA SSVGKPSLIN TGRHVNEMAE
401 EGFPVDLLS NFVNTKETVD KTEKLINSEE YMRONIDHIN RLSKWKESA
451 SNTENVEFGN ENA

11AA_SEQUENCE 1.0
PI:E83776 - hypothetical protein BH1013 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: E83776
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirama, C.; Nakamura, Y.; Ogasawara, N.; Kuhara, S.; Horikoshi, K.
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis.
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: E83776
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-155 <STO>
A:Cross-references: GB:AP001510; GB:BA000004; NID:910173440; PIDN:BBB04732.1; GSPDB:GN00137
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1013
C:Superfamily: Bacillus subtilis hypothetical protein yfjM

E83776 Length: 155 May 30, 2002 09:26 Type: P Check: 7616 ..

1 MNTLCIIPCG NKKIMQKQED IGPVASEAY IGTILRLCAM YAEKFDHNV
51 ILSAKHGFLE PDDVNDGPYD VFSFHSVGV ITFDQLTQV HEKLLDSVPH
101 VVLLTGKKYR PIVEACPEEA TVECPLLSYS GIGYMQALK RAIDKOPPLH
151 SAKRN

11AA_SEQUENCE 1.0
PI:S55948 - hypothetical protein YLR392c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein L8084.i3
C:Species: Saccharomyces cerevisiae
C>Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 29-Oct-1999
C:Accession: S55948; C36445
R:Du, Z.
Submitted to the EMBL Data Library, January 1995
A:Description: The sequence of S. cerevisiae cosmid 8084.
A:Reference number: S55944
A:Accession: S55948
A:Molecule type: DNA
A:Residues: 1-518 <DUZ>
A:Cross-references: EMBL:U019729; NID:9625097; PIDN:AAB82352.1; PID:9625110; MIPS:YLR392c

A:Experimental source: strain 5288C (AB972)
R:Ackerman, S.H.; Tzagoloff, A.
J. Biol. Chem. 265, 9952-9959, 1990
A:Title: ATP10, a yeast nuclear gene required for the assembly of the mitochondrial F₁-F₀ complex.
A:Reference number: A36445; MUID:90277691
A:Accession: C36445
A:Molecule type: DNA
A:Residues: 1-210 <ACK>
A:Cross-references: GB:J05463; NID:91431794; PIDN:AAB05631.1; PID:91480201
A:Experimental source: strain D273-10B/Al
C:Genetics:
A:Map position: 12R

S55948 Length: 518 May 30, 2002 09:26 Type: P Check: 3922 ..

1 MAPKISISLN PRYNGEYSS NDMGSIYSL QUTKALSTIK ISYLKGESE
51 TLTKIDQYEM FOONGMMMPG QDNKSEPTLM KEOQVFPED NWNALDGSS
101 KPEKVPGSY NYSQFDKFP RKPECLKNHT AKTAVFVTRS NARLPPEFNS
151 HMQEPNKTIDN LDLYFYSFGK VIYMOVQLE LCKSSSWKRP FHLIRELET
201 FEETPEPRDL IIEPEDDNE ELNAPSNNSR GNSWVTNNEF FNSSLKVPSS
251 KDQKVVNCGV YKSDRNFSQ ANSILLENQ INSRVSVVT STROSTRLVN
301 GKMYFPSTYK MGLPDGESNM RIEVRSRDLK QYRRDYLFY SGSONFDKRY
351 VMEGNIASL SKMOITPLKL QNLILETTY LSQGIANGNY SSLKLEIDL
401 NQKSNKPLD DLNIRENPD GSMFECELRL KHPILRLKY FNEEDYRHNG
451 NRLYSFKTCT IKRFTSIQLL IEMGINKRK QSEVNDIPQV IFCQVREHVE
501 AEALPRVVP PTYTEMAS

11AA_SEQUENCE 1.0
PI:F82557 - hypothetical protein XF2449 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82557
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing and Analysis, Sao Paulo, Brazil.
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: F82557
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-544 <SLI>
A:Cross-references: GB:AE004053; GB:AE003849; NID:9107631; PIDN:AAE85248.1; GSPDB:GN00128; XFSC:XF2449
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvaranga, R.; Alves, L.M.C.; Araya, J.E.; Baia, G.S.; Baptista, C.S.; Barros, M.H.; Bonaccorsi, E.D.; Bordin, S.; Bove, J.M.; Britones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H.; Colauto, N.B.; Colombo, C.; Costa, F.F.; Costa, M.C.R.; Costa-Neto, C.M.; Coutinho, L.L.; Cristofani, M.; Dias-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
Submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme, M.; Fullan, L.R.; Garnier, M.; Goldman, G.H.; Goldman, M.H.S.; Gomes, S.L.; Gruber, A.; Ho, P.L.; Hohelsel, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigret, F.; Lambais, M.R.; Leite, L.C.C.; Lemos, E.G.M.; Lemos, M.V.F.; Lopes, S.A.; Lopes, C.R.; Machado, J.A.; Machado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.A.L.
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.;

Miyaki, C.Y.; Monteliro-Vitorello, C.B.; Moon, D.H.; Nagai, M.A.; Nascimento, A.L.T.O.; Netto, L.E.S.; Nhai Jr., A.; Nobrega, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Paris, A.; Peixoto, B.R.; Pereira, G.A.G.; Pereira Jr., H.A.; Pesquero, J.B.; Quaggio, R.B.; Roberto, P.G.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki, H.E.
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira, J.F.; Silvestri, M.L.Z.; Siqueira, W.J.; de Souza, A.A.; de Souza, A.P.; Terenzi, M.F.; Truffi, D.; Tsai, S.M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zago, M.A.; Zatz, M.; Meidanis, J.; Setubal, J.C.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2449

F82557 Length: 544 May 30, 2002 09:26 Type: P Check: 144 ..

1 MICRRVYLV SMNLRSSH FVPLLSVAL LLOGCKDSP SAKROLIMVE
51 AKLPPLPLK DIPPLFLDIE RTFOFFWDT TNEVNGITPD RPSRPFASI
101 ASVGFALTAY PIGENGWIS RNOAIDRTLT TLKFLRDPM GPQRTGCAGY
151 KGFYHFLDM QHGRYDSWV ELSSVDFTALL MMGLVFESY YDGEDVREKE
201 IRIADKLIR RYDWTYLOQR KPLISMWYP ERGFIEHDM GIVEAMVYL
251 LALGSPTHPL EPVSWDEWTR TYNKDMGVFQ GOEYLAEGPL FGHQYTHWV
301 DFRIDIDQWY RERGIDYFLN SRAVLAHRD YAIIDNPKMK DYGENVGLT
351 ASGPQNTQY EYRGQRQFF HRSKAGLGF EFPDDGTIAP TATVASIVTA
401 PEVYIPATIE MKRYGDFLY SSYGFDAFN PSEDIYVPLK TGRILPGRGV
451 VASDYIGIDQ GILLMIANY RNDPFWNMK KMKYVPTGLE RAGFIGWILQ
501 EDEVLAGPQK DERAAVRSI GIAESHAPPA QTORVPSLSG KKEE

!!AA_SEQUENCE 1.0
P1:F82112 - hypothetical protein VC2147 [Imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: F82112
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Umayam, L.A.; Gill, S.R.; Nelson, K.E.; Read, T.D.; Tettelin, H.; Richardson, D.; Ermolaeva, M.D.; Yamahayan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; McDonald, L.; Utterback, T.; Fleischmann, R.D.; Niernan, W.C.; White, O.; Salzberg, S.L.; Smith, H.O.; Colwell, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833
A:Accession: F82112
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-118 <HEI>
A:Cross-references: GB:AE004287; GB:AE003852; NID:g9656689; PIDN:AAF95292.1; GSPDB:GN00126; TIGR:VC2147
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2147
A:Map position: 1

F82112 Length: 118 May 30, 2002 09:26 Type: P Check: 2237 ..

1 MKNTTLPSE QTQGEALKIA KATQRPQTK EQTKLITQGI EKGIALYKKQ
51 OKKHRAQDK LRKALKAKO SSTIEIHAD DYAAELGDDS HINQAKLAWV

101 LIALSWIGFI SYLWONS

!!AA_SEQUENCE 1.0
P1:C84898 - hypothetical protein At2g46060 [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84898
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487
A:Accession: C84898
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-787 <STO>
A:Cross-references: GB:AE002093; NID:g3702340; PIDN:AAC62897.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g46060
A:Map position: 2

C84898 Length: 787 May 30, 2002 09:26 Type: P Check: 3197 ..

1 MAELIFLLR TLVVFSCALS IONCLODAAD NNGNEFYVSS FRPESVYRP
51 YDRIYINVDL PWFSSLNVA IESVDITAK SISKISSSL PVICFRDGP
101 PLDPASTNAL QGLELGRFFN GSFERADPSE IAQCYPQOK NITRLTNEQ
151 CLFVLILSY NIDTAQIVRS SAFSPSANIS VEGCKTATMW GPFCNQTIYP
201 LSCRFPNQOT ASYISCADSF PSSCLTGAET KYVALDVDCI AEOLVIMASN
251 VKYDSNESYL MCYARFENA SETLHDYAAD IHKPLYLVNK PKAGRMYIVI
301 SLSSRENKFA QGNNSSSRVC PSINNVKVLGC PVGACAPRCG QOYIYLQAVM
351 RRGMLTPPQOS YFPVVDASL GGSSTNPFLE PIVSNFSSIP ELDTSWTYF
401 LMNIPQGGSG GHIFRRLSD STIOYEVYLR FGLPTIDDR DYYVVRTSA
451 SRSMFFELYN SSKEMVDFTI LYAREGTSF GLRDLDSNT PAASRGSPTL
501 VSLSLERCPR GGSYGCORF AFDAAGLTSY RFTYLESAS ICKHDSNFCG
551 CDRTHGFDC SIEIVSHQER IVQSIALIAS NAAALLPAYW ALQRREYPM
601 VLFSSGSSS ALYHACDVGW WCVLSYVNLQ FMDWLSFMA VGTVPYVLSF
651 AGEAVKRTIH TVVAITLALL ALTOATFRASN IIVLAIGSL GLLIGFLVER
701 VTKYRSYCGS AGFSLNMLDR PRAVKEWFSN LITLTKRRR WGFVAAGLVA
751 FTMAAISFKI AIVNHENLH NGADNYELTR ODSLSRN

!!AA_SEQUENCE 1.0
P1:C84582 - hypothetical protein At2g19880 [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84582
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.

Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A:Reference number: A84420; MID:20083487
A:Accession: C84582
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-550 <STO>
A:Cross-references: GB:AE002093; MID:93687230; PIDN:MAC62128.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g19880
A:Map position: 2

C84582 Length: 550 May 30, 2002 09:26 Type: P Check: 6852 ..

1 MSLDSDAI LPSLRATFS PPAVVOQOG CTICLLALG WLALEYVRNR
51 EVRRKNSIK AGNSIAFLYO DINELEHSRO VKLPVSVVM PLKGFGEHNL
101 HNNRSQITSL YGGPLEFLFY VESTEDPAVH AVSRLSLMYQ DHVEAKYVVA
151 GLSTGSOIKI HNOLIGVEKA HKOTKYVFL DDDVRLHGT IGAITTEMK
201 NPEVPCSMGF ATGGRTFFELM GGCMMAHADD FRDPRYGVVS GLRDGYSDD
251 MTLASLAGAH KRLITSPVA VEPHPLASDL SFGRYWNYLR KQTFVLESYI
301 SKYMWIMINKA LPAVHCYLSM GFVAHYVMAI IHITSALRIY IKGYHOLEDT
351 TSASGGLSVF LMLHIDRRMR FYVGMVLVIT LAICTFIELL SMNNLRREV
401 QLCNMLSPFA PRSLATYVM GLPFIIGHI SPMYRGFNKP RTGVFNVED
451 YPYPNLXPHT IFONIKSALE NKHFVETPL LLOTVSSSIW LSTSLPRGI
501 LKRPSTSSA ASWPETMYK GLNGTSLYL PARAIQNIHI SNIAWSARWC

!!AA_SEQUENCE 1.0
P1:AD0772 - hypothetical protein STY2348 [Imported] - *Salmonella enterica*
subsp. *enterica* serovar Typhl (strain CT18)
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhl
A:Note: this species has also been called *Salmonella typhl*
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AD0772
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, C.; Mungall, K.L.; Bentley, S.D.; Holden, M.T.G.; Seakula, M.; Baker, S.; Basham, D.; Brooks, K.; Chillingworth, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, K.M.; Dowd, L.; White, N.; Farrar, J.; Felwell, T.; Hamlin, N.; Haque, A.; Hien, T.T.; Holroyd, S.; Jagels, K.; Krogh, A.; Larsen, T.S.; Leather, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quall, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.G.
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhl CT18.
A:Reference number: AB0502; PMID:11677608
A:Accession: AD0772
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-703 <PAR>
A:Cross-references: GB:AL51382; PIDN:CAD02498.1; PID:916503362; GSPDB:GN00176
C:Genetics:
A:Gene: STY2348

AD0772 Length: 703 May 30, 2002 09:26 Type: P Check: 9075 ..

1 MEODRLLNA MPEMCNKNP INDGQREWHI ADIFGLREE RYDELDERYN
51 QALTESFTSR EAEKRYEFAM NOMDNPEYDM DTIVEAGPQG LALIKNQRA
101 RPRSTHAWLA ENQYWNHRAW LYRSIGWARE TTRAMWICAA ACNERAWYIA
151 LNAIDCEPRQ WMAAALTSTN SKVFGQPMWL VEPLEGADVA GQPLMEDLAE

!!AA_SEQUENCE 1.0
P1:T20228 - hypothetical protein C54G10.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T20228
R:Matthews, L.
Submitted to the EMBL Data Library, June 1996
A:Reference number: Z19240
A:Accession: T20228
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-660 <WIL>
A:Cross-references: EMBL:Z75532; PIDN:CAA99810.1; GSPDB:GN00023; CESP:C54G10.3
A:Experimental source: clone C54G10
C:Genetics:
A:Gene: CESP:C54G10.3
A:Map position: 5
A:Inserts: 86/3; 133/2; 305/3; 336/2; 442/1; 628/3
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

T20228 Length: 660 May 30, 2002 09:26 Type: P Check: 7424 ..

1 MARLADFLF FRRFRLLKI LEPCHRRSIT AMLAGALLIA TALDOVMTYL
51 VGIOPSLEFYV ALGNKADTF KILCARGAAI ILGKAFTLAV FKYLTNMLAI
101 KSRQVCNLTM HRLYFKRQAF FKLSGSDML DNPQDRLOTD IEKATRIISN
151 DLLAPIATAP FIIGYTYWLT YESSGWIGPA AIYYFCIQT IINKMILSPI
201 VQKVSDEKMA EGDFFORHME VRSNVEAIAF YRAGVENIM TNQKLKMLIE
251 TQKSLTEWRM VLSITNVPD YFGGILSYLI IGVPFIYTHL YDDVSPAEIN
301 GIVSRNAFFY LYLYFSSTV LKLTGDFEL AGVTHRWEL HEELIRLHSD
351 CLETFDRPSTP VPSGVVVIAS DEDDKSASRN MOEIKGOMS LERDEQEEEE
401 AQYLGGKGTG QEDDWPDDGV AITVDSATIS PRNDHSILIV QLSLQIIIG
451 QTLITGDSG CGKSSLLRMF AGIMHCSSK MDCHRRILTS NLFLAOKPY
501 FPSGNTTLRQ QIYYPVKALQ VDKDVARITQ ILEWVKMEL VERCGGLDTP
551 VEMDMKTLTS PGELQRLSLA RVFYTKPRIV FLDESTSAIG FELEMAIYRK
601 LQEKIITVS IGHRIYSLQF HDMELRVKGR SGENSLHID TASIASRTAS
651 FLAGDTVLNM

11AA_SEQUENCE 1.0
PI:H87334 - conserved hypothetical protein CC0691 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: H87334
R:Niemman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.F.; Alley, M.; Ohta, N.; Maddock, J.R.; Potocka, I.; Nelson, W.C.; Newton, A.; Stephens, C.; Phadke, N.D.; Ely, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolonay, J.F.; Smit, J.; Craven, M.; Khouri, H.; Shetty, K.; Utterback, T.; Tran, K.; Wolf, A.; Vamthayan, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87334
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-208 <STO>
A:Cross-References: GB:AE005673; NID:g13421912; PIDN:AAK22676.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0691

H87334 Length: 208 May 30, 2002 09:26 Type: P Check: 7467 ..

1 MPEIIVNCP MLNLTADILQ GRVYRLPEPT VDHHDLEKAA IDCDPASWEI
51 MSVNGCGEG EDWFGALQGE TDRGERIGFA IRLVDGKV GTSYINIR
101 LHGGLEIGAT FLNPARSRP VNPSKRLML GHAFDRKAI RBLVTDVKN
151 ARSQAAIOLK GATKCGVLIN HKVTWGHVR DTAVFSTIDY DMPAIRERLE
201 FRLESTFA

11AA_SEQUENCE 1.0
PI:E86254 - hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E86254
R:Theologis, A.; Eckert, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altar, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.T.; Koo, H.L.; Kremetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, A.; Miltischer, J.; Miranda, M.; Nguyen, M.; Niemman, W.C.; Osborne, B.I.; Pal, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.; Tortum, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vayberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: E86254
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-783 <STO>
A:Cross-References: GB:AE005172; NID:g3157932; PIDN:AAJ17615.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

E86254 Length: 783 May 30, 2002 09:26 Type: P Check: 9544 ..

1 MATLGDIGA AAINILTAII FLAFAIARI QPNDRYFP KYLKGIRSS

51 PLSGALVSK EVNNGISYL RLNMMPAL KMEPELIDH AGDSAVYLR
101 IYILGKIFV PIALAMSTL VPVNTSHQL QLAIRNVT SDDIKLSISN
151 IENGSDSLF GRFWHLVMA YAFETWCYV LMKYEVAA MRALFQNEQ
201 RRPDQFTNLG LSOLLSQVLY RNVPADPDES ISDSVEHFEL VNHDPYLRH
251 QVYVANDLA ALVEOKKSTO NMLDYQLKY TRNDRKPRI KGFGLAMCK
301 KYADIDHYA EIEKLINEQR KVKKDDPSV MPAAFVSFKT RMGAASAGT
351 QQSDPLEML TEMAPAREV FMSNALPIYV SLVYRLIMH IAFETLTFE
401 MIPAFVQSL ASTEGIEKNA PLKSIIBND LFKSVIOGFL PGIVLKLPLI
451 FLPSILMWS KEGGFVSLSS LERRAIFYR IFNLINFLG SVYTSAPFQ
501 LDSEFLKQSAK EIPKYGVAVI PIKATFFITY IMDGNAGIA GELIRKPLI
551 FPHKNSLV KTEKDREEM NGQINVHAT EPRIQLYELL GLVYAFVTPV
601 LLPFIIFFA LAVLYVRHQI INVYQEVES AARWPDVHG RIISALIIAQ
651 ILLMGLISTK GAAQSTPELL FLPIITFFEH RYCKGRTEPA FLRHPLEAM
701 VKDTERARE PNFNLKPYLQ KAYIHPVKD NDYEDSRFDE ISGCTIEDSD
751 EECVYVPTKR QSRINTPAVS HASRGSRSR PSK

11AA_SEQUENCE 1.0
PI:C96782 - unknown protein F22H5.13 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96782
R:Theologis, A.; Eckert, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altar, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.T.; Koo, H.L.; Kremetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, A.; Miltischer, J.; Miranda, M.; Nguyen, M.; Niemman, W.C.; Osborne, B.I.; Pal, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.; Tortum, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vayberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: C96782
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-654 <STO>
A:Cross-References: GB:AE005173; NID:g10022280; PIDN:AAJ12693.1; GSPDB:GN00141
C:Genetics:
A:Gene: F22H5.13
A:Map position: 1

C96782 Length: 654 May 30, 2002 09:26 Type: P Check: 7509 ..

1 MSTTSYRVR LAFVALLSAT TFYCIHKRR LKHKNLSLN PSSYLKASNG
51 KIFPISOTGT AKALQRLHE LCASNDIAFD IVDPHSTPEE DLPKETLVAF
101 IASTWDGPKR PKNGEFLVM LGESAEDERV GSLLSDCKF AVFGVSSRAY
151 GESYNNVAKE LSRMIGLGG LEMIPVGBD VDGDGLDRAF QDWCGVITV

201 LKGGSAOETN GVSQIGAVE DDLEYDSTD EEDENDADG GIVDLEADIAG
 251 KAPSKRNGVY KATKYDCKE WTPVIRASL TKOHDEYLLQ GYKIGSHSG
 301 VKICRWKQSO LKRGCGCYK SFYGIESHRC METTPSLACA NKCFECRRHH
 351 TNPYKSMOW KMDPSVIYK GALLDHLKMMI KOMKGVPCVT PEKLOEGLNP
 401 RHCAISLJVE PIMYPEINAL VDELHGRIS TELVTNAQFP EKILMMKPIT
 451 QLYVSVAAT KESLKAIDR LEADWERI DSLKALOQK QRYVRLTLV
 501 KGMTEELDA YENLFSIGRP DEIEKGVY CGSSATSKLT MENVPHTDV
 551 KAFSEALSLK SNCEYVACE HAHSCCVLLG RTEKPKVDGK WFTWIDYEKF
 601 HDLVASEGFP TSTWDYMAQTP SMAVYGAQGC GFDRGQLAYK KERHHPKPO
 651 AVLA

IIAA_SEQUENCE 1.0
 PI:A8488 - hypothetical protein At2g45230 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: A8488
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.;
 Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblum, T.V.;
 Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.;
 Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.;
 Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.;
 Copenhagen, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg,
 S.L.; Fraser, C.M.; Venter, J.C.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: A8488
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1374 <STO>
 A:Cross-references: GB:AE002093; NID:g2583130; PIDN:AAB82639.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g45230
 A:Map position: 2

A8488 Length: 1374 May 30, 2002 09:26 Type: P Check: 8740 ..

1 MRISMNCQG VGNTPTVYRHL REIRGLYPE VIFLCETKR RNYLENVVG
 51 LGFDLHIVE PIKSGGIAL MKKDSVOIKV LOSDKRLIDA LLIMQDREFY
 101 LITCIYGEVQ AERGLMERL TRIGLSRSGP WMLTGDFFML VDPEKIGCP
 151 ARKSSCLIEF RQMLNSCGLW EVNHSQYFS WIGNNDELV QCRLDRIVAN
 201 QAMMEPEQA KATYLOKICS DHSPLINLV GDMWRKMGF KYDRWVORE
 251 GFKULCWFV SQOSTKTNAL MMEKIASCHR EISKWKRVSK PSSAVRIQEL
 301 QFKLDAATQ IPPDRRELAR LKELSQEYN NEEQWQESK RIMMRNGDR
 351 NTKYFHATK NRRANRIQK LIDEGREWY SDEDLGRVAE AYFKLEPASE
 401 DVGTVEELE NLTPVSDQM NNMLLAPITK BEVQRATFSI NPHKCPGPDG
 451 MNGELYQOFW ETMGDQITEM VQAFRRSGSI BEGANKTNIC LIPIYILAER
 501 MTDPRPISLC NVIYKYVIGKL MNRKLKIIP SLISETQAF VKGRILSDNI
 551 LIAHELHAL SSNNKSEEF IAKTDISKA YORVEMPFLF KAMRGGLPAD
 601 HWIRLMCEV KSVRYOVLIN GPDPHEIIPS KGLRGDPLS PYLVICTEM

651 LVKMLQSAEQ KNOITGLKVA RGAPISHL FADDSMFYCK UNDEALQIIT
 701 RIIEEYSLAS GQRYNYLKS IYFKHISEE RCLVYKRLG IERBGEGVY
 751 LGPESFGGS KVALSTLKD RLCKKVLGMO SNPLSPGCKE ILLKAVAMAL
 801 PTTWNSCFKI PTKIQOIES VMAEFWKKK KEGRLHMKR WCHLSRPKAV
 851 GLGPKKEIEA FNIALGKQL WRMITKDSL MAKVKSRYE SKSDPLNAPL
 901 GSRPSFARKS IYEQVILIKQ GIRAIVNGE TINWTDPMI GAKAKAKAQA
 951 VKRSHLVQY AANSIHVYKD LLIPDGRDM WNLVSLLEPD NTOENIILAR
 1001 PGKRTDRDF TWEXSRSGHY SVKSGYWVMT EIIQNRNPD EVLQPSLDPI
 1051 FOQTKKLDPV PKIHFLMRC VNNCISVSN IAYRHLARK SCVCPSPHGE
 1101 TVNHLLEKCP FARLTWAI SP LPAPGGEMA ESLFRMHV LSVKSGPEE
 1151 SDHALLIPMI LMRLMKNRD LVFKGREFTA PQVILKATED MDAMNNRKEP
 1201 QPQVTSSTRD RCYKQPPSH GWYKNTDCA WSKDLGNGCV GWLERNHIGR
 1251 LMLGLRALP SQQSVLETEV EALRMVLSL SRFNTRYVIF ESDQYLVSL
 1301 IQNENDIPL APRIQDIRNL LRHFEVYKFO FTRREGNNVA DRTARESLSL
 1351 MNYDKMTSI TPDWIKMLVD LETV

IIAA_SEQUENCE 1.0
 PI:A8448 - hypothetical protein At2g03400 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: A8448
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.;
 Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblum, T.V.;
 Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.;
 Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.;
 Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.;
 Copenhagen, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg,
 S.L.; Fraser, C.M.; Venter, J.C.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: A8448
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-131 <STO>
 A:Cross-references: GB:AE002093; NID:g4335759; PIDN:ADU17436.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g03400
 A:Map position: 2

A8448 Length: 131 May 30, 2002 09:26 Type: P Check: 6617 ..

1 MWQOSLSLTL TIGGSVYKSS LLNRNLNSVK ASSLIGDRCV SCQFLRKSPS
 51 FRSHKSLKQ RNLIRVEARW PFQGGEGGL DPSSERSESA NEDILIFFQ
 101 IDLATRYOVK TKCRVFISFK MVANGFDVLQ Y

IIAA_SEQUENCE 1.0
 PI:H83554 - hypothetical protein PA0736 [imported] - Pseudomonas aeruginosa
 (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: H8354
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey,
 M.J.; Brinkman, F.S.L.; Hufnagle, W.O.; Kowalik, D.J.; Lagrou, M.; Garber,
 R.L.; Goltzy, L.; Tolentino, E.; Westbrock-Wadman, S.; Yuan, Y.; Brody, L.L.;
 Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, R.; Lim, R.M.; Smith, K.A.;

Spencer, D.H.; Wong, G.K.S.; Wu, Z.; Paulsen, I.T.; Reizer, J.; Saier, M.H.; Hancock, R.E.W.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen.
 A:Reference number: A82950; MUID:20437337
 A:Accession: H83554
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-358 <STO>
 A:Cross-references: GB:AE004508; GB:AE004091; NID:9946611; PIDN:AMG04125.1; GSPDB:GN00131; PASP:PA0736
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA0736

H83554 Length: 358 May 30, 2002 09:26 Type: P Check: 7512 ..

1 MPIDRQAQ QRAVDIVFQ RELQRDQDQ AFRLPQAQ QLAHYHALL

51 DSTRDRPDID HDLRSQLSL GNRVASFGA LAFASVFL FYRWGLFPT

101 VAQVAILGS APAEFAFLM VOAKDASGYF SKLAAMVAF CFVLDITMLG

151 QIFNVPSDL ALVPALVAL LLAYLCNARL LLAAILCYM GFIAAVGTW

201 GGGYWLVSVE RPENFPPAA LIFAVPLCFD QRNFSGRAVI YRVFALLGLF

251 LPMVLVANWG SGGSYALPSA LIEGLYVAG FVAALVITWL GARRNADVS

301 NTGITFEVIE LYTKFFDWMW EAMPKYLEFL VLGSLALLIL LVLRRLRTP

351 GIAARDA

!!AA_SEQUENCE 1.0
 PI:F83173 - outer membrane protein OprC PA3790 [imported] - *Pseudomonas aeruginosa* (strain PA01)
 C:Species: *Pseudomonas aeruginosa*
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: F83173
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brinkman, F.S.L.; Hufnagle, W.O.; Kowalik, D.J.; Lagrou, M.; Garber, R.L.; Goltzy, L.; Tolentino, E.; Westbrook-Wadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, R.M.; Smith, K.A.; Spencer, D.H.; Wong, G.K.S.; Wu, Z.; Paulsen, I.T.; Reizer, J.; Saier, M.H.; Hancock, R.E.W.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen.
 A:Reference number: A82950; MUID:20437337
 A:Accession: F83173
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-723 <STO>
 A:Cross-references: GB:AE004797; GB:AE004091; NID:9949950; PIDN:AMG07177.1; GSPDB:GN00131; PASP:PA3790
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: OprC; PA3790

F83173 Length: 723 May 30, 2002 09:26 Type: P Check: 9154 ..

1 MEKRMSTQQR AAGNACPTAA FSPDPAQLAQ RRRWAGFAA LGLALSPSA

51 LLAERSHQH DHAELAPSV VTGAQSSPL TIVTNPKEPR QVPASDGAD

101 YLKTIGFAV IRNGSGNDP VLKGFSGRL NILTNGMML GACPNDMDAP

151 TSTISEPTY KLTVIKGPOT VLKPGASAG TILFERPER FGLGSRVNA

201 SLIAGSNGR DKVLDAAGN RLGLYLRFTGN HAOSDDYEDG AGNTVPSRMK

251 KMGDAVAVG TPDEDTLIEL TAGKDGEAR YAGKMGDSQ FRRESLGLRF

!!AA_SEQUENCE 1.0
 PI:F83610 - probable fatty acid desaturase PA0286 [imported] - *Pseudomonas aeruginosa* (strain PA01)
 C:Species: *Pseudomonas aeruginosa*
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: F83610
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brinkman, F.S.L.; Hufnagle, W.O.; Kowalik, D.J.; Lagrou, M.; Garber, R.L.; Goltzy, L.; Tolentino, E.; Westbrook-Wadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, R.M.; Smith, K.A.; Spencer, D.H.; Wong, G.K.S.; Wu, Z.; Paulsen, I.T.; Reizer, J.; Saier, M.H.; Hancock, R.E.W.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen.
 A:Reference number: A82950; MUID:20437337
 A:Accession: F83610
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-392 <STO>
 A:Cross-references: GB:AE004466; GB:AE004091; NID:9946120; PIDN:AMG03675.1; GSPDB:GN00131; PASP:PA0286
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA0286

F83610 Length: 392 May 30, 2002 09:26 Type: P Check: 1486 ..

1 MWFNGFDLS PMQLAVTVL LTHITIVSYT VYLIRYSNHR ADDLHRLQH

51 FFRWMLWTT GANTREMTAI HKHHAKCET VEDHSPVHK GLFTVLRAGA

101 ELYKAEAKNQ DTLRIYKNC PDDWIERNLV SRPIGGVVL MALIDLALFG

151 ALGLTVAVQ MWVIFPMAG VVNGIGHAVG YRNECRDAA TNLVPCIL

201 GGEELHNNH TYNSAKLSV KKWEDLGA WIKFSLRL ARVAVAPIA

251 HRYEGKSLD MDTAMILNN RPOIMAYRK LVIAPLVKOE VAKADESVRH

301 LFRRAKRLS RETSLQDRH HVRIDSMALH SQALKVIER RLALQDTMAK

351 TSANGHMLA AIDWVHEAE NSGIOSKEF AAQIKTYSLR PA

!!AA_SEQUENCE 1.0
 PI:S73657 - MG288 homolog F04_orf260V - *Mycoplasma pneumoniae* (strain ATCC 29342)
 C:Species: *Mycoplasma pneumoniae*
 A:Variety: ATCC 29342
 C>Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
 C:Accession: S73657
 R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
 Nucleic Acids Res. 24, 4420-4449, 1996
 A:Title: Complete sequence analysis of the genome of the bacterium *Mycoplasma*

pneumoniae.
 A:Reference number: S73327; MUID:97105885
 C:Accession: S73657
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-260 <HIM>
 A:Cross-references: EMBL:AE000031; GB:U00089; NID:g1674002; PIDN:AAB95979.1;
 PID:g1674006
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library,
 November 1996
 C:Genetics:
 A:Genetic code: SGC3
 A:Start codon: GTG

S73657 Length: 260 May 30, 2002 09:26 Type: P Check: 3241 ..

1 MQAFREYWK OKKDYTDKKQ LLEALKLSFA KEONKTRAFI IKNFQGISN
 51 YYPNDEEDS EAKTAFGTQ GIAFPQSLK GIFMSEWLAK OLGEKAKINL
 101 DIKSLAVTDS KISPTIKMK DIGIKRNDK PYNRFEDI EYOGNYKLSW
 151 LEATIAKFSG IPEWKGKLN LKFIYDGLS WEIYQKPDYP GSLFQDDOK
 201 OQLLFRLHW EKITVOEPEF MELIKSONLH NLELRTESTK PPVVDLASYL
 251 HYQLKLNDQ

11AA_SEQUENCE 1.0
 PI:E69365 - hypothetical protein AF0925 - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: E69365
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, R.J.; Gilm, M.; Hickey, E.K.; Peterson, J.D.; Richardson, D.L.; Kertavag, A.R.; Graham, D.E.; Kyrpides, N.C.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kinkness, E.F.; Dougherty, B.A.; McKenny, K.; Adams, M.D.; Loftus, B.; Peterson, S.; Reich, C.I.; McNeil, L.R.; Badger, J.H.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Uitterback, T.; Cotton, M.D.; Spriggs, T.; Arlach, P.; Kalne, B.P.; Sykes, S.M.; Sadov, P.W.; D'Andrea, K.P.; Bowman, C.; Fujii, C.; Garland, S.A.; Mason, T.M.; Olsen, G.J.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon Archaeoglobus fulgidus.
 A:Reference number: A69250; MUID:98049343
 A:Accession: E69365
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-315 <KLE>
 A:Cross-references: GB:AE001040; GB:AE000782; NID:g2689363; PIDN:AAB90325.1;
 PID:g2649684; TIGR:AF0925

E69365 Length: 315 May 30, 2002 09:26 Type: P Check: 7306 ..

1 MLESMTSSL ILFVALVI SLILYCSYL RRERMKISG IFSVSAFAY
 51 ILLEFSFVN LLSNLLFILL VELAVFILL PIKVVPELK FEMBEKYEI
 101 DGIPVIGYE KKKKYNAFY TPKRKIFVT KSLKDLIGE ELKAVIYHES
 151 GHSKNKMMI TRSTAMFVY LIAAVLTTL FLEMGKFOF NLKVSFRTL
 201 GALLIYATF FVFWFINIH EADLFVAVKS GYENFSALF KITFYVNLGD
 251 YAEVQKIDL KNFNSGVTP FEILKILIKQ SIYLFPPNI LNQIPQTHP
 301 PLRYRILLAH QTLKC

11AA_SEQUENCE 1.0
 PI:T01809 - hypothetical protein A_TM021B04.3 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 A:Reference number: 214248
 A:Accession: T01803
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-3449 <KAP>

C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 22-Oct-1999
 C:Accession: T01809
 R:Dante, M.; Wamsley, P.; Gibson, A.
 Submitted to the EMBL Data Library, June 1997
 A:Description: The sequence of A. thaliana TM021B04.
 A:Reference number: Z14440
 A:Accession: T01809
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-938 <DAN>
 A:Cross-references: EMBL:AF007271; NID:g2191181; PID:g2191183; GSPDB:GN00063;
 ATSP:A_TM021B04.3
 C:Genetics:
 A:Gene: ATSP:A_TM021B04.3
 A:Map position: 5
 A:Introns: 125/1; 559/3; 664/1; 798/3

T01809 Length: 938 May 30, 2002 09:26 Type: P Check: 5621 ..

1 MKRENVLVS LLEFVIVFLM QVGEAQRIT NVNVIYVNDI GTAVSNMTLL
 51 CINSLSDFY SSHPETQTRL VTTVDSKND VTTAAAGTR IPISLNHAL
 101 RYVCVINHQF ICTLLIYIA FSSKALDLIT NKEVKAILGP WTSMAQFMI
 151 EMGQSQVPI VTSATSPSL ASIRQYFPR ATYDSSQVH AIKEIIRLFG
 201 WREVAHYVD DTFEGEIMPR LFDVLQELNV RIRPRVISP NARDELISVE
 251 LRRMTLPTR VEVVHLVELL ASRFAKATE IGLKKQGYW ILTNTITDVL
 301 SINNETELET MQGVLGKTY VPSKELENF RSRWTKRPEI SDLVNYGLMA
 351 YDATTALALA IEAGTSNLT FYKMDAKRV SELQIGYGSQ YGPKILQTL
 401 RYRQGLAGD FQINELQD SYFEIYNWG QGGRTIGWM KEYGLEKAVD
 451 QPKASKTFS SMQDRLEPI WPGDTTSYPR GMEIPTNGKR LQIGVPVNT
 501 FQGVKATRD PLNSTIFSG FSIDYEAVI QAIYDLSYD FIFPDGQYD
 551 ALYQYVYLG YDAVVDATT SSNSMYDF SLPTPSGVG LVVPKDSVR
 601 RSTTFIMPL TLALMISL SFFIIGLVW VLEHRVNDP DGPQYOLST
 651 IFWFSFIMV FAPREVLSR WAWVVIWY FLVIVLQSY TASLASLTF
 701 QHLHPVTNI NSLLAGESY GYQSSFIQR LRDGSESEAS LVSYGSEHC
 751 DALLSKQAE GYSAVLMFY PYRIFLGQY CNKYKMQTP FKYDGLGFV
 801 PIGSPVADI SRAILVEES NKANOLENM FKPIDESCPD PLNPNDRPS
 851 VSFROLGFD FVVLIVMAI VCTMALIKV YQFLKENDQ RNLRVIMEKF
 901 NEPDQSYIK DYTCCOSSG QGMPKNGOG ANAVNNGN

11AA_SEQUENCE 1.0
 PI:T01803 - hypothetical protein T10P11.5 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 14-May-1999
 C:Accession: T01803
 R:Kaplan, N.; Johnson, D.; Schutz, K.; Guoj, L.; Hoffman, J.; Tili, S.; de la Bastide, M.; Granat, S.; Hamed, A.; Gottesman, T.; Hasegawa, A.; Shohdy, N.; Parnell, L.; Dedhia, N.; Johnson, A.F.; Lodhi, M.; Martienssen, R.; Chen, E.Y.; Wilson, R.; McCombie, W.R.
 Submitted to the EMBL Data Library, November 1998
 A:Description: Sequence of A. thaliana BAC T10P11 from chromosome IV.
 A:Reference number: 214248
 A:Accession: T01803
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-3449 <KAP>

A:Cross-references: EMBL:AC002330; NID:g2262135; PID:g3892055
 A:Experimental source: cultivar Columbia
 C:Genetics:
 A:Map position: 4
 A:Introns: 62/2; 86/3; 115/2; 138/3; 157/3; 198/3; 262/3; 284/3; 323/1; 726/3;
 960/1; 994/2; 1221/3; 1316/1; 1868/2; 2858/2; 2961/2; 3063/3; 3112/3; 3427/3
 A:Note: T10P1.5

T01083 Length: 3449 May 30, 2002 09:26 Type: P Check: 8548 ..

1 MKWGTLLKDL KDKVGAETT ADLIGBAIS DPTTPSSSQ ASPSSSPAL
 51 AQHDFULLSP TSROKLKLEL DFKRYWEEFR SSSSEOEKEA ALNLSVNTFC
 101 RLVKOHAND QLVTLVLEPH IFSFVIGRAF VADVERLKDG SSHGANLLTA
 151 IEVLASGPPD KOSLDSGIL CCLITFNFAF LTVSASEGE KTVNVEEKVE
 201 GSVVNTMKAL ASHPSAOSL IEDDSLQLLF KMYANGSLMA FSRFVGLVS
 251 FHNIOLEKNA MOILGLLVN DNGSTASYIR KHHLIKVLAM AVKDEPDCG
 301 DSAFTVGIYD LLECEVELSY RPEFGVRLK DDIRNAHYH FLYOFALILS
 351 SMPKDIYFAP DHSPHKNRG SNDCKOPPL SLKTRONDS EKOQSLILNS
 401 RONDEFALH FSPALSRLLD VLVTLAOTGP IESSGTSTSI LSQRLTGYS
 451 RROTPSANRR YDEPCEQSG KVKDLEAVQM IODIFLKAEN KDLQAEVLNR
 501 MKRIFSHLE NYRICOELKT VPLLVLMMG FPSILOELL KLEAVAVTV
 551 NCVPBELLIS LCELIQPID SELKHTLSF FVKLTSFDOQ YKRVLGEGV
 601 LEVLODDLKO HKLLRGPDQY SGVSNHLDYV PSSPSFKOHL DSOQAIISSP
 651 KIMESGSGKL PFEVERITIT VGMDCMISL KNSOVNOEAF RBAAGTVTL
 701 PFLIDENHT SILRIFSLI TGDIKQVHHE ELEALIDVAK SGMVTRVSGD
 751 QYKLHYEVC DIMGALMRIV GVNGSAORVF GEATGFSLIL TLHTFQGE
 801 BCRDESHLV YIKLFKHLR LITTAVCENA INRMKLSHY TSQTYDYLLY
 851 EGGLECVBLE RHVIOLELE ALEVLPPL TSESMASEM ABCEKASFLY
 901 KTASGOFPND KOKIYNAGAV RVLIRSLILC TPRLQLEFLN ILERLARASP
 951 FKKETITSAG CVELLEITIY PFIQSSPFL SHALKIVEVL GAYRLSPSEL
 1001 KMLCRIVMOM RVNANGPSLI GMMKELIME EDTGLECVSL APFVEMDSK
 1051 TGHASVQVSL GERSWPPAAG YSFVCWVQFR NELTQELLES EYKAGASSK
 1101 TPLISGOQSE QNIFRIFSVN AISNGSPSYA ELVFOEDGIL TLATNSNSL
 1151 SFSGLETEEG KWHHLAVVHS KPNALAGLFQ ASVAIYIDG KIRHMGKAGY
 1201 SPSPYKSLQ VIGTSATCA RACGDSMAI LDLIDTMSG GIOKFEEDNR
 1251 QGDSNAHSG IWMDLRLGN LSTQLPGKL IPAFDQCSG FPARAGSPSL
 1301 VLVUDPLSA ASLIGGIPRF GRUVGNVSLC RONVIGNSIR PVGMAVVIA
 1351 LVEAESRDM LHMALSLIAC ALHONSONVK DMETTYGYHL LALFLRPKA
 1401 LFDMLCLEIF FOISACEAF SEPKKLESQ TTISMSPTEL IPENNYEDPT
 1451 LCKFOETISS VSGHGDMDF SGRKDSFSL SELEMGNPV ETSNCIVLSN
 1501 ADMVEHLID WTLWTAIVS IOIASLGFLE NLISILMYS HMLILRQIN
 1551 LKHLHLVTLQ RGDVAVLVE KIVILLRCLL ENGFLEPELE DVYRAIMTF

1601 NPPEIKSONS SMRESMKHV IVRNLVLEML IDLQVTIKAE ELLEQWHKTV
 1651 SKRLITYFLD GAVHPSMHR IMTLIGVCLT SSFNESLKEF ASGGYQGLVR
 1701 VLOSFDSPD IYIIFCLIF GKPYPRLE VRALDRHALM PDDGSVEIN
 1751 FVDLDSVVA MAKSTFDRLI MOSMLAQSG NLSQVSARCV ABLVGYADM
 1801 TGEIQKALM HKTYARLMG GEASAPATAT SVIRFMYDLA KMCPOSANAC
 1851 KNEPIQKCA DLYFCVRAF HAVKIAQLS MKAEONITG GDDSSVEGNF
 1901 CRVSHODMST KTSISAGSFP QDQTSVIVSY DMVTPSDYVA VDKVENFLT
 1951 PGESNKSFO GREYIAKODG DHVGSVSASS EMKSLDTGS SSQOVPIDSR
 2001 SRSFPMLES PLSEKSSLE VPFIIPSFSKS STISTRPSH ISVSEPDASS
 2051 DQSSGSQSS AVHTLFTISP KVLLETDESG YGGGPCSAGA SAVLDFMAEV
 2101 CADIMTEQIK AVQALESILE MLPYVDEPC VVVFQGLCLS RVNNTLERRE
 2151 LRDEEDDDK LDKRKRSANL DAFQMIVDR VYKGAFOPT GVLRTLEPLL
 2201 SILQLANKDG RVEEYTSOGK GLLSIGRATR QLDAYVHSIL KNTNRTIYLC
 2251 FLPSFLITIG EEDLPSRLGL LVESTKKOTS KLSGREGID VSAVLQTLVA
 2301 NKNIILCPNS LDTDLNCCIC VNLSILHDQ RKNVQNMASN IIRYLLVHRK
 2351 SALEDLLVKK PHRGKFDVL HGFDRLLTG NLPEFSKMLE SSEQITTKVL
 2401 EOGAAVMWID YIAGSAKEPD VRMGKMDGR TREMGRLRD TSKLDKHE
 2451 QVNERRYALE VVRDMSAEL RVRONKTYGL ILHAESWPT HIQOLVHEG
 2501 IFFMRISHGV EDLKNOLCPI EGPYRMRRKL ERCKLKIDSL HNLBCKLEL
 2551 GEIELLSKKS EDGLVSDMD SEPAFLSEL YSEFSSEAD DLKDVPSARN
 2601 GWNNDATST NAASLHNSLS FCGKSSSTAV SVPISVATDE KSETSPITS
 2651 SSGKMEIKH VEESEKELK DDGEYLIRPY LEHLEKIRFR YNCERVAGID
 2701 KHDGIFLIGE ICLVYIENFY IDDHGICIEK ECEDELSIID QAQGLKQPH
 2751 GLEESKSKS TLMSTTIKIG AVGGRAMAYG GCAWGKEKVA VTGNLPHPMH
 2801 MMKLDVHET LKRDYELRRV AVEIFSMDCG NDLLVFHKKE REEVFRNLIA
 2851 MNLPRNSMD TTISGAKOE SKESGRLEKL MAKSTFRMQ NGEISNFQYL
 2901 MHLNTLAGRG YSDLOIYPIF PWILADYGE SLDLSDPNNF RKLDRPMGQ
 2951 TPEGEFEFR RYESWDDEPV POHYGSHYS SAGIVLFYLI RLPPSAENO
 3001 KLOGGOFDHA DRLFNSIRET WLSAAGKNT SDVKELIPEF FYMPELENR
 3051 FNLIDGEEKS GDKVGVILP PMARGSVREF IRKHREALES DVSINLHHW
 3101 IDLIFGHKOR GKAENAVNV FYHTYRGNV DVDAVTDPM KASILAQINH
 3151 FGOTPKQLOF KPHVKRRTRD KVPHPPLKHS MHLVPRIRK CASSINOIIT
 3201 FNDKLLLTGA NCLLPRGYK KIIRMGFPDR TLFNMGYDQ KLLSTHENH
 3251 EGNQIQCAGV SHDGRIYVTG AEDGLVSVWR VSKDGPGRSR RULREKSLCA
 3301 HPAKVICLRV SQPYMMIASS SDDCTVIIMD LSSLSEVRQL PNFSAVTVY
 3351 YINDLTGEIV TAAGSVLAWV SINGDCLSV NTSQLEPDLI VSAASTSD

3401 WLETTWYVTG HOSGALKVWR MVHCTDPIKC CYRMQKQOSE TESRRFDNN
 I:AA_SEQUENCE 1.0
 P1:T48284 - hypothetical protein T22P1.220 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48284
 R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.;
 Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z24490
 A:Accession: T48284
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-428 <BEV>
 A:Cross-references: EMBL:AL162971
 A:Experimental source: cultivar Columbia; BAC clone T22P1
 C:Genetics:
 A:Map position: 5
 A:Note: T22P1.220

T48284 Length: 428 May 30, 2002 09:26 Type: P Check: 3770 ..
 1 MTLPFLAAV EVLQLSTLT VAEIKSETIS NDSRPVILE KFGIIEIGHV
 51 TVSVSVSVL SPILDSKLG FEVLSESLP HVLLEQONF SFCVLDSHYI
 101 LHFTTFDLS PPPRSQESKS YPITSNDPS LFEANCPEP RVSKKVTETI
 151 YHDIYPNGSR DYLLAGSAQL PGILYVEFLC YLSFLCFMFC PCWNHKOIVK
 201 RIHLMTALL LVKSLFLICA AVYKHVYKT GAHGMNIVF YIFQFISVL
 251 LFWIVVLGN GMSFLPKRLH VKEKKLVIV VPLQVLANIA SIVGEGPY
 301 TDQVNSMNOI FFLADITCCC AIVPAMWSM CCLRETSKTG GKAVKNLAKL
 351 PVLRKFYVL IGYLFTFRIV VVVMKKKADF TYQWVSNAE EIATLSFYCL
 401 MFWMRPIEK NEYCDVDEE EIVELSLK
 I:AA_SEQUENCE 1.0
 P1:T39938 - hypothetical protein SPBC23E6.04c - fission yeast
 (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T39938
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Pohl, T.
 submitted to the EMBL Data Library, May 1998
 A:Reference number: Z21892
 A:Accession: T39938
 A:Status: preliminary; translated from GR/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-1649 <MOO>
 A:Cross-references: EMBL:AL023287; PIDN:CAI18872.1; GSPDB:GND0067;
 SPDB:SPBC23E6.04c
 A:Experimental source: strain 972h-; cosmid c23E6
 C:Genetics:
 A:Gene: SPDB:SPBC23E6.04c
 A:Map position: 2

T39938 Length: 1649 May 30, 2002 09:26 Type: P Check: 5030 ..
 1 MASSLQKQK NISNNVLKI NKIRRAPSL YDPKVAADM LELIVTVAS
 51 GFHELAHVP RLTYFEKTL GDSVOVDYR LNRTEKETI DLECYQILRL
 101 LAPFTTERNA LKYLEMLIR FSIHEVVSDE FILSELPFD HPFARIIGC
 151 SKPKSRPLF LENAIMPVT LSRADIVHAL SRDKKEFAAF AQPONTAES
 201 HNMPELARF WAGTMEVLY AMHSSNEDPN VLLDRFFLAV SYANVSYSI

251 DEQIAGFML SSIAASDPLS PSIIPLVSA ITRLSFONI KPALICGHL
 301 LQFCSPFED HEQLEKLESF GASSLIELS QEHRDEFFV SYWVSLIKSR
 351 KOKDKRLIS LDTISQIR VTHEQAKFL SVIPNODK ALQSYRLD
 401 SVIOPERREG KLDNLINTLQ DKRSSTFSK KDEVILKRI SEIDSQTSFE
 451 QCLAVADSAA DLDSVFISL LSKFGDKIP LFCJANGSE RIILSLIEL
 501 KRTIEKNQV DYQITLVLV YSIQSKDTEY NSRALNLIIT FLELRNMLE
 551 FSIYGMNDN DNKNLRMLSP VETKYYCSDI LIDRSSEIGL DGYLFSYIP
 601 ERLFEKKPK NASKEIAVTS FLSSHAQSK LSNVAVILLE ILFVHGKVE
 651 DAKMOILPR LEQISENSE KFTVSKREY EALVNCFNHT SFTSLSLSL
 701 SNIVLSQAIC RRIEIQSHL KDPQRLFVK AVISODEOPH YYVDVLSIK
 751 IPDVFFKKLI GSVRLVKEKN PATAKKRRID SHIPGDVQR LTRILELLET
 801 KNAASYPKLA SPLFVLNSV IALKEDIVSS NYLLOLLIGL LYEMIGASPI
 851 TELSPSIRID TLVGCIRSTN NPOIQNKAL LVSAIANAP EAVLHGVPPI
 901 PTFMGSTVLS RDAFSLHVI EQTVKTVISA LTRIGKDFDS SLVSCFPNA
 951 PPHIQHRRL RLYRLVLOTI GSNRFLSVL IOPAEKMLLA KSTNVVAIHD
 1001 FCLTLVQSFV VADRIGSINQ CSRFCLSLE EQSNSDSNGK AVSLIKIDEL
 1051 PMDVOLATIG SLRVKLELI SLVSKAKNFA FDLAKIMENS VDSEVEIQAG
 1101 LFESTIKLIT LSOQSSNEME LGHYVVALRS VHLHPNLEF CTVLGKLLHD
 1151 ERALLRRRAL SIYQQRVQOG SKVSALTALI PDVYNISNY SDEETTQIAM
 1201 DCLAVMAKRF SASPELFI SP IEVYSGPYGL KNSARDVQS AIVCTIVLTN
 1251 TLAARILPEL ADIVYISLTI LDDARKDPEG DLELACFSM MIDPFKVLPE
 1301 FSSSYVEPTI KCALASDRAF EHDAGIGELP ETIANEIPPR LLMKSIFAM
 1351 PECARLGSTFA ALRLLELIEL ALQNSSRSAL GYVYSIFNF FLDSRDSRHS
 1401 LLFAEDVDNV ETQAVNVPFLK FVKKLSDTTF RPLFLHLHSM ALLEDIYETDP
 1451 SGIVSRQTFE YNFLTIFLDT LKSIYVNYA YVLDITIEL SSKDTNSEVR
 1501 HLVNSSLVSA FENDTEEFMM VPARFGKISP VLIHQIQVAP LLDKVLVKA
 1551 IVELASVASS SDNFRSMNTQ LLOYLRSSNI NARLAIQIQ TOLYGRIGEN
 1601 WISTLPQSVF FIAELMEDDD DOYETATAEL VRIIDRLIGE NESIQDYILT
 I:AA_SEQUENCE 1.0
 P1:S36335 - U2 snRNP 40K protein - Trypanosoma brucei
 C:Species: Trypanosoma brucei
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
 C:Accession: S36335; S28637
 R:Cross, M.; Wieland, B.; Palfl, Z.; Guenzl, A.; Roethlisberger, U.; Lahm,
 H.W.; Binderl, A.
 EMBO J. 12, 1239-1248, 1993
 A:Title: The trans-spliceosomal U2 snRNP protein 40K of Trypanosoma brucei:
 cloning and analysis of functional domains reveals homology to a mammalian
 snRNP protein.
 A:Reference number: S36335; MUID:93209224
 A:Accession: S36335
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-326 <CRO>
 A:Cross-references: EMBL:X69934; NID:g10563; PID:g10564

A>Note: part of this sequence was confirmed by protein sequencing

S36335 Length: 326 May 30, 2002 09:26 Type: P Check: 3173 ..
 1 MRLTLTIR APOFNALRO RELDIRGLI TVLEHTLTF LNDSPDVLNL
 51 SQNPRLRLR FPGDSAPLAT AAQNSSAKP ASRMMLRLQT LVYHRRRLTH
 101 VSEATCAAVL PNLRAVADH NEFRRLDRL FLSHKKLEI LSIENPITI
 151 SEDNARLRAV VVELCPTEKL VNYQRVQVD RQNVETMRKE FVGLVGMRR
 201 LEKKQLLOQ SAPTEADASA SESVKKIKRR SRHAREASK NSADTAPRP
 251 AASKVETPAT PSVAADAGEE EASNAIQAR LEALEEKMAA AETEEELMEL
 301 QDELTELETL MKHQASKQT KKTRTS

11AA_SEQUENCE 1.0

PI:T19297 - hypothetical protein C15A7.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T19297

R:McMurtry, A.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z19104

A:Accession: T19297

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-814 <WIL>

A:Cross-references: EMBL:267735; PIDN:CAA91530.1; GSPDB:GN00028; CESP:C15A7.1

A:Experimental source: clone C15A7

C:Genetics:

A:Gene: CESP:C15A7.1

A:Map position: X

A:Introns: 34/3; 84/3; 191/2; 243/2; 279/2; 325/1; 396/1; 423/2; 476/1; 548/2; 611/2; 663/2; 698/2; 744/1

T19297 Length: 814 May 30, 2002 09:26 Type: P Check: 7776 ..

1 MFFPLMLVL FNTIPIRVSE ADESESRHOA YLDHRIKIED VLGPLDGSMP
 51 SELPVGIGAO LKLDHISRYD EKTGFLTLAL TVKFMWSDER LQMDPSHYSG
 101 FPRTRLEFVD FWLYQIMKPR IYMNITERR QSKTIDLLSN TLVENDIHSK
 151 GFWMTTKIL LKTFCELDER GYPHPONCS FSEIPNNAN DIFLGSMUK
 201 AKRTDFTDP QQVVRVQDFQ INDVTTEDLY LFMNRIYLD LTRPRSLVR
 251 SPVSPFIVIQ RKNVYVHLQI YAPMIFICTC LIISGFLPSE VALPLLVENL
 301 ISEYLYEHNI KDVLPTDFDG TPITALATF TLAETMSIIG WKMFITITMT
 351 KKQKAYELNS SKFPLDSGPF NIRVLRRLTL IDKFLWMLI AOSFGPTTI
 401 FLOLFEPQQA ELRLMLTLKVN RQNPPLVAS QYSVKINISF SLSEESYVNS
 451 YFHCARYPS SCGYTMNGK KLHGFDDDHG FMSVINAEL VSYLYQPYE
 501 SFIFQWIDE RLKMPNAYS GVRIVEKTF EFRKDNNCWM PLVKRSTYR
 551 RSELDLLEF SDARTLISTK GEIKTALQTM VTKQPSFG EXPNDYONCS
 601 IMILPNQAD EFRVSPNGY CNPKFENLEH RAVRVHDLH MGVESNIEFY
 651 FANTYTAEE VTGEFYMAKT EFRNLMFKR VNKLLNWKLS ISILISMEL
 701 ILAGLFPNGY SIFGYAYCFE VEVHGLYVS KILPNDIRGI PYHGALALCF
 751 LIETVFLFCW KVFYSYARNO KLFAQLPNVF ADNRVEVLKE VYVVRVVCV
 801 ALCLQPLSVI YKTR

11AA_SEQUENCE 1.0

PI:T23772 - hypothetical protein M162.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T23772

R:Baslam, V.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19797

A:Accession: T23772

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-170 <WIL>

A:Cross-references: EMBL:282278; PIDN:CAB05254.1; GSPDB:GN00023; CESP:M162.2

A:Experimental source: clone M162

C:Genetics:

A:Gene: CESP:M162.2

A:Map position: 5

A:Introns: 75/1

T23772 Length: 170 May 30, 2002 09:26 Type: P Check: 6239 ..

1 MFKILFLAI LIPSSLTDCP DIDGTFREIK GICKKFTSL MKYEDARDMC
 51 HHNNPAGPSW LAVVPDQETS NFLAVYAGSI FGEGFKHFWI GLNRDPISKT
 101 LSWDTGLSVS YTNFGSNVAQ NYFSENTINT KMNTLGDEEV HNFVCSYRPS
 151 TVPATVTRQP QARRLAAMKN

11AA_SEQUENCE 1.0

PI:T26443 - hypothetical protein Y113G7B.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26443

R:Lenard, N

submitted to the EMBL Data Library, September 1999

A:Reference number: Z20215

A:Accession: T26443

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-331 <WIL>

A:Cross-references: EMBL:AL110477; NID:e1542121; PIDN:CAB54331.1; CESP:Y113G7B.9

A:Experimental source: clone Y113G7B

C:Genetics:

A:Gene: CESP:Y113G7B.9

A:Introns: 32/2; 58/1; 178/3; 212/3; 247/3; 285/2

T26443 Length: 331 May 30, 2002 09:26 Type: P Check: 4393 ..

1 MEITIDLLCL LGITSAIITI LMINLVYKI VLNPSKRKND WYLYYRFTL
 51 DIFGAGLFS YIAYTLNME APEFWQYRS LIYVLAIPWS HISTCSIIA
 101 LSIISIRISA TCEPIIYFKN RKKIPNMPVL LIGSLGLAE EYWLSEFGCY
 151 NMEIPKTCIV FGCATNOCPE HWMLIQRSII FSLIVPSLI LSIKILMMS
 201 VKHQOSNNQI SKANRLALD TCTVLLPDL PARCGHMPPT APWFSPNNG
 251 SYNPVLKITG CAIESIVYTR VLLFRTPENS ASTPVOKLIY TLRFSLSLAQ
 301 FPDIFKNIFE HFLRSRAVTV LYKNILKLIQ F

11AA_SEQUENCE 1.0

PI:H83904 - hypothetical protein BH2040 [imported] - Bacillus halodurans

(strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: H83904

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirama, C.; Nakamura, Y.; Ogasawara, N.; Kuwara, S.; Horikoshi, K.

Nucleic Acids Res. 28, 4317-4331, 2000

A>Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*.
 A:Reference number: AB3650; MUID:20512582; PMID:11058132
 A:Accession: H83904
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-818 <STO>
 A:Cross-references: GB:AP001514; GB:BA000004; NID:g10174613; PIDN:BAB05759.1; GSPDB:GN00137
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH2040

H83904 Length: 818 May 30, 2002 09:26 Type: P Check: 9494 ..

1 MTRKFINELK NDDLDGEMC FFPNHLISA GFLPKWLTQF TFPDYMALVE
 51 EKRRLEKWN ETISECNNEI EKMLDLTKQ KKKLISOLRR YKIPYAKGT
 101 ELLANIGAGF QIKLEDMIEK MNYLEAQISK TFQTELDQR HMKQVYLEN
 151 ESIOEAVFLS SPDAVHNGLK HFLSKELBOS RPRGQRQEQ LAAYIDQRT
 201 SKNETASYG PTNYGTFTSQ PGLLELHVG PIKROLEMSY WAVQSLANM
 251 AEDDSILPYL KPKLSPFIKR ENSNLRSSN GKLIOLEPIY HKMIRYSDGN
 301 HNHIQIAML DLSTACIAR IOKLKEKRII LLEIDFPTSV FDPPEKLEK
 351 IQOLPNDCCS KSKMLHIVNE WSSHLLTWKE SESFOQRRII LTLLENSFEO
 401 NVGLSRKRD GKHYDRLAY FEARGDVET CRIGLNMOK QKDOLEPIFK
 451 LISRAVEN RALTEVAYKE YQSFNSDVNF LTFALHMGK KTELQYAKE
 501 KLEEDNRLS EWTALHFEK GSIYAKEDV DAFCLNPPQ DVALLSPDIM
 551 IAADVDVAIN EKKYKLIGE VHSGLQWMSV LNSVYDQLR LNEEYHHLG
 601 PTLOSFLER VCPRAPKFT RELSOGTIV ENLGRSMKSR EYVSTAEIG
 651 LIYENERFVY TEGEKPLMD LETDVEPLNQ IFSLPYSKSF SIQIGHTPR
 701 IEINGVYFOR ERWTFACKEL LYRVNGYTDW LLLDMATDLR KNYMPHYVY
 751 ARGNNNEKPI FVDFHNFPL EYLYOLKRN EHVSTENVP AODSLMFTRG
 801 ENKHTAEFRF SVIHENLT

11AA_SEQUENCE 1.0
 P1:BA4135 - RNA polymerase Ecf-type sigma factor BH3882 [imported] - *Bacillus halodurans* (strain C-125)
 C:Species: *Bacillus halodurans*
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: BA4135
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirama, C.; Nakamura, Y.; Ogasawara, N.; Kuhara, S.; Horikoshi, K.
 Nucleic Acids Res. 28, 4317-4331, 2000
 A>Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*.
 A:Reference number: AB3650; MUID:20512582; PMID:11058132
 A:Accession: BA4135
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-154 <STO>
 A:Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BA07601.1; GSPDB:GN00137
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH3882

BA4135 Length: 154 May 30, 2002 09:26 Type: P Check: 5308 ..

1 MKVSEIYEEL KODLYRFART ISRNEQAND LIQDAMEKSL RCELLDMPR
 51 HKQKAMPFRV MKQOLDIRR KKKRETDMD SVIPPTTAG TSYEMWDL
 101 SKLPASISDL VKRHFWLGLT SQEISEKGV PASTVRYRIR LAVKRLREYM
 151 KEEG

11AA_SEQUENCE 1.0
 P1:D97337 - mismatch repair protein Muts-like ATPase [imported] - *Clostridium acetobutylicum*
 C:Species: *Clostridium acetobutylicum*
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C:Accession: D97337
 R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, H.M.; Dubois, J.; Qiu, D.; Hittl, J.; Wolf, Y.I.; Tatusov, R.L.; Sabathe, F.; Doucette-Stamm, L.; Soucaille, P.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium acetobutylicum*.
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: D97337
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-577 <KUR>
 A:Cross-references: GB:AE001437; PIDN:AAK81487.1; PID:g15026659; GSPDB:GN00168
 A:Experimental source: *Clostridium acetobutylicum* ATCC624
 C:Genetics:
 A:Gene: CAC3563

D97337 Length: 577 May 30, 2002 09:26 Type: P Check: 4438 ..

1 MDEKLYWYS AIAFLALGIS AFVFPOLAIG YNNIYIFLGT ALSTAAIFLA
 51 FYMKAKIIRK LNTFRFRYW GEFQKRRNT RNVRKFEFEH KEDNADEFNV
 101 DEQFWMELNM NKVPELADR LTPSGEMLY KIFKPFESK EKLERNII
 151 KIFQNNKEVR EKVGLLEIKL GRKKENGVD LMKQLEVDY KYKVFANFL
 201 WGLIASILTI PIKGYIIL LAFLIMNTI EHNKERNKVE LYVOSLGLN
 251 GIINTANRIS KINCPERKY TDSLKATSSK LMKVAKKTAG IEREGVDII
 301 GDALYNIPI EERKFNSIN DIRKLKELK ELTKALGED ALMSTASRQ
 351 WIQYCEPEF VDNRIIKS EYHPLLENP VNSIDLDGK GIMLTGTMA
 401 GKSTFLRTIG LNSLQAOTY TCAKYYKTS FFKIMTSISP EDNISGKSY
 451 YFREAEALKR IINQSDRRP VLCIIDEIR GNPPIERVVA SAEILNITEK
 501 HNTLTVATH DLELTELKE DYLCYFSED IDDEGLQFDY KLYGICKTR
 551 NAVKLKYLE YPNEIIQKTN DRLAKIL

11AA_SEQUENCE 1.0
 P1:A10020 - probable membrane receptor protein YP00164 [imported] - *Yersinia pestis* (strain CO92)
 C:Species: *Yersinia pestis*
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
 C:Accession: A10020
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; Sebatina, M.; James, K.D.; Churcher, C.; Mungall, K.L.; Baker, S.; Basham, D.; Bentley, S.D.; Brooks, K.D.; Cerdeno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Fellwell, T.; Hamlin, N.; Holroyd, S.; Jagsels, K.; Leather, S.; Karlyshev, A.V.; Moule, S.; Oyston, P.C.F.; Owall, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.G.
 Nature 413, 523-527, 2001
 A>Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: A10020
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-686 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC89027.1; PID:g15978269; GSPDB:GN00175
 C:Genetics:
 A:Gene: YPO0164

A10020 Length: 686 May 30, 2002 09:26 Type: P Check: 7936 ..

1 MISGKNKVK CKLLACAVTS AISAFASNAI ASNAIADTTS IAAINRSAIN
 51 ISSHNHNSKK LNDADVIYTV APLYSPIVWV TSPKVPQPV PASDGDIDYLR
 101 TIFGFSLVRN GGTNGDVER GMEGSRILKIL TDGAELIGTC PSRMDAPTST
 151 ISPESEDLIT VVKGPSVLW GPVGSAGTIL FERERPHFDQ SGVKGNA5VL
 201 AGSNGRMDKN LDASLGNNEG YLRMMGNQSR AADYKKGNNV RVSQMDKN
 251 ADMALGWTPT TDTLELSVG KNGEARYGG RGMDSQFLR ESLAARYKS
 301 NIGEVLDKIE AKINTYVDH IMDNKTILRT PMAMSSNLD RMLGGRVWG
 351 TWLMQDLKLD AGDMQTNTH RKNKGNDK NARMNNGVF GELTMAASEQ
 401 NKLSGARLD RHQVNTYTA GEPTRSATLP AAFMRVEHNL ANAPVMLYAG
 451 LGHTEREPPD WELFSPKFGP AGSTSAFEGV KSEKTOIDI GGOYSGEQR
 501 GWSAVYIGRV NDFLEFKYDP KNLRTSQADN INATIMGEM GMEYALTPYW
 551 KADASLAIYAW GENTRDNRPL POIPPLEARL GLTFEAGDWS SSALMRLVNS
 601 QHRVAINEGN VVKGRDSSA GFGVLSANVA YKVNKQVKLS SGIDNLLINKA
 651 YSEHLNLAGN SGFGYSVNTP INEPGRTLMA KLNLTPE
 11AA_SEQUENCE 1.0
 P1:AB1288 - transport proteins homolog lmo1706 [Imported] - *Listeria*
monocytogenes (strain EGD-e)
 C:Species: *Listeria monocytogenes*
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AB1288
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.;
 Bloeker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve,
 E.; de Daruvar, A.; Denoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.;
 Durand, L.; Dussurget, O.; Ertlan, K.D.; Fsihl, H.; Garcia-Del Portillo, F.;
 Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.;
 Jackson, D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.;
 Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordisiek, G.; Novella, S.; de Pablos,
 B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes,
 N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland, J.; Cossart, P.
 A:Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AB1288
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-289 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAC99784.1; PID:g16411160; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo1706

AB1288 Length: 289 May 30, 2002 09:26 Type: P Check: 1598 ..

1 MMKKVYKVK QNGIYOYGOA VSARYGRNDV SGNAQIAYV MLESTPMLL
 51 IAAITLALYLH IDKDSYFNMI KEFAPQDQMD FLEENLNL TLTKNGGILSTI
 101 GIATILMSAS NGMNAVVKSL NKAYGVNTKR NYVQRLLSM FFLTAMLATV

151 GATILLLVG QQIGMPLINH LNFSEDFLSF WNNLRMTVTL IYIFVETPL
 201 YWVAPNRST LISVLPGLF STIGMTVASV GFAYVNNFG NYSATYSGISG
 251 VIITLMLWF LTGIIIMIGG ELNATLAIK KKEIGEIN

11AA_SEQUENCE 1.0
 P1:A11659 - transport protein homolog lln1818 [Imported] - *Listeria innocua*
 (strain Cl1p11262)
 C:Species: *Listeria innocua*
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C:Accession: A11659
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.;
 Bloeker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve,
 E.; de Daruvar, A.; Denoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.;
 Durand, L.; Dussurget, O.; Ertlan, K.D.; Fsihl, H.; Garcia-Del Portillo, F.;
 Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.;
 Jackson, D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.;
 Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordisiek, G.; Novella, S.; de Pablos,
 B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes,
 N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland, J.; Cossart, P.
 A:Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: A11659
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-289 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CAC97049.1; PID:g16414320; GSPDB:GN00178
 A:Experimental source: strain Cl1p11262
 C:Genetics:
 A:Gene: lln1818

A11659 Length: 289 May 30, 2002 09:26 Type: P Check: 1480 ..

1 MMKKVYKIK HNGIYOYGOV VSARYGRNDV SGNAQIAYV MLESTPMLL
 51 IAAITLALYLH IDKDSYFNMI KEFAPQDQMD FLEENLNL TLTKNGGILSTI
 101 GIATILMSAS NGMNAVVKSL NKAYGVNTKR NYVQRLLSM FFLTAMLATV
 151 GATILLLVG QQIGMPLINH LNFSEDFLSF WNNLRMTVTL VYIFVETPL
 201 YWVAPNRST LISVLPGLF STIGMTIASV GFAYVNNFG NYSATYSGISG
 251 VIITLMLWF LTGIIIMIGG ELNATLAIK KKEIGEIN

11AA_SEQUENCE 1.0
 P1:AC2420 - hypothetical protein alr4915 [Imported] - *Anabaena* sp. (strain PCC
 7120)
 C:Species: *Anabaena* sp.
 A:Note: *Anabaena* sp. (strain PCC 7120) is a synonym of *Nostoc* sp. strain PCC
 7120
 C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
 C:Accession: AC2420
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.;
 Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.;
 Matsunoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.;
 Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing
 Cyanobacterium *Anabaena* sp. strain PCC 7120.
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AC2420
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-313 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA076614.1; PID:g17134053; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr4915

AC2420 Length: 313 May 30, 2002 09:26 Type: P Check: 9610 ..

1 MKRLDYCCGS FSQLNVSSSR KRGNQYKPI LVLTVIDLIT QGLITNOIR

51 VSEELIQTFE RYWSLIGSES YKGLHYPPE HLQSDGFVHL VKEPFENGIO

101 PKTKNLKEA VEVAYLDKEL FNFLQDESPP QELIDALVTT FFOEOODELE

151 EMQIQNSQ DADLATETSS GSINDNNPK WGCRAIIRN APERKIVRY

201 YDYKAPCGL RYTKAINONI VDGAIKPPA QFYDSRIHNG IALCKHHMA

251 FDRGWFTVDE QYKIIVSKEL QEISPHSKPM KDFHGERILL PNOEQYLPFL

301 ESLQWHRQWY FQA

11AA_SEQUENCE 1.0

PI:AF2463 - CAB/ELIP/HLIP superfamily of protein [imported] - Anabaena sp. (strain PCC 7120)

C:Species: Anabaena sp.

A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002

C:Accession: AF2463

R:Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shlimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120.

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AF2463

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-59 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAW6961.1; PID:g17134401; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: asf5262

AF2463 Length: 59 May 30, 2002 09:26 Type: P Check: 4158 ..

1 MPTDKISAS VVEDRNSWRW GETPOAEIWN GRLAMIGFLA ATLIEFFSG

51 GFLHFWCITL

11AA_SEQUENCE 1.0

PI:AG2271 - CAB/ELIP/HLIP superfamily [imported] - Anabaena sp. (strain PCC 7120)

C:Species: Anabaena sp.

A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002

C:Accession: AG2271

R:Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shlimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120.

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AG2271

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-59 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAW75425.1; PID:g17132860; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: asf13726

AG2271 Length: 59 May 30, 2002 09:26 Type: P Check: 4679 ..

1 MSGFKNPKST VTESPNAVRF GETSESESWN GRFLAMIGFLS IYLIEAFSGQ

51 GFLHFWCITL

11AA_SEQUENCE 1.0

PI:AH1862 - CAB/ELIP/HLIP-related protein [imported] - Anabaena sp. (strain PCC 7120)

C:Species: Anabaena sp.

A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002

C:Accession: AH1862

R:Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shlimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120.

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AH1862

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-40 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAW72407.1; PID:g17129794; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: asf10449

AH1862 Length: 40 May 30, 2002 09:26 Type: P Check: 580 ..

1 MGFHQSESW NGRFLAMIGFL AAIAIEFFSG QGFLHFWIIL

This Page Blank (uspio)

1	337: EHRLD	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(Y)WXX(L) EFFVSYWVSL	IKSRK
1	YL92_YEAST	ck: 3922 len: 518 i P18634 saccharomyces cerevisiae (baker's ye (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(S)(H)WXX(F) PTFNSHQOE	NKIDN
1	146: NARLP	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(R)(F)WXX(L) ESFYRFWAGL	AYDRI
1	YXIK_BACSV	ck: 7683 len: 153 i P42302 bacillus subtilis. hypothetical prot (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(R)(F)WXX(L) ESFYRFWAGL	AYDRI
1	Q9X4B6	ck: 2369 len: 504 i Q9x4b6 escherichia coli. k30 capsule biosynt (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(S)(F)WXX(F) QSFSSFWDGF	TGHDN
1	293: GEGRP	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(S)(F)WXX(F) QSFSSFWDGF	TGHDN
1	Q9L1E1	ck: 7894 len: 70 i Q9L1e1 streptomyces coelicolor. sc3d11.21 f (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(E)(F)WXX(L) LSDFEFWAHL	NK
1	59: DGDQ	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(E)(F)WXX(L) LSDFEFWAHL	NK
1	P95529	ck: 7071 len: 492 i P95529 pseudomonas putida. tdng protein (ec (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(R)(Y)WXX(L) EEFIRYWMHL	RSSEW
1	449: REAFG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(R)(Y)WXX(L) EEFIRYWMHL	RSSEW
1	Q44310	ck: 4158 len: 59 i Q44310 anabaena sp. (strain pcc 7120). rna (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(H)(F)WXX(L) QGFLHFWGIL	
1	50: ELFSG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(H)(F)WXX(L) QGFLHFWGIL	
1	Q52209	ck: 5380 len: 324 i Q52209 pseudomonas putida. pupr protein. 12 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(R)(Y)WXX(L) QHETRYWVTL	KPRMA
1	310: LOLEV	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(R)(Y)WXX(L) QHETRYWVTL	KPRMA
1	Q9F0G8	ck: 6799 len: 364 i Q9f0g8 rhizobium meliloti (sinorhizobium me (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(F)WXX(L) EAFADFWSKL	SLAFG
1	144: EAVPY	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(F)WXX(L) EAFADFWSKL	SLAFG
1	Q06954	ck: 8871 len: 460 i Q06954 vibrio cholerae. rfbg protein. 6/200 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(F)WXX(F) EKFEVFWDP	VFEYN
1	158: HLLGG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(F)WXX(F) EKFEVFWDP	VFEYN
1	Q00620	ck: 5873 len: 683 i Q00620 pseudomonas stutzeri (pseudomonas pe (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(D)(Y)WXX(F) PLEKDFWQRF	LDSLK
1	462: GLGHT	QRFPDYWELE	SGGAD
1	Q930E9	ck: 4054 len: 567 i Q93qe9 azotobacter vinelandii. phb synth (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(S)(F)WXX(Q) PSFTSFWSAQ	TPFVA
1	5: MDOA	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(S)(F)WXX(Q) PSFTSFWSAQ	TPFVA
1	Q93N64	ck: 559 len: 314 i Q93n64 coxiella burnetii. hypothetical 3 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(Y)WXX(L) LDFVDYWFGL	RDSNY
1	164: LCCHL	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(Y)WXX(L) LDFVDYWFGL	RDSNY
1	Q9P981	ck: 4445 len: 385 i Q9p981 robilliardea sp. (strain y-20). end (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(E)(F)WXX(L) EOPGFEWGL	ASREN
1	150: TAATT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(E)(F)WXX(L) EOPGFEWGL	ASREN
1	Q9UW21	ck: 7245 len: 429 i Q9uw21 candida albicans (yeast). hypothe (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(C)(F)WXX(L) EYFTCTWDEL	PNNKS
1	128: NPVLG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(C)(F)WXX(L) EYFTCTWDEL	PNNKS
1	Q9U123	ck: 6679 len: 190 i Q9u123 homo sapiens (human). pro0529. 3/ (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (C)XFX(C)(F)WXX(L) CHPVCWFVCL	PLRVS
1	18: VIVCV	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (C)XFX(C)(F)WXX(L) CHPVCWFVCL	PLRVS
1	Q9NZH3	ck: 2970 len: 280 i Q9nzh3 homo sapiens (human). c-type lect (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(Y)WXX(L) EFFYSYMTGL	LRPDS
1	191: SOSYS	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(Y)WXX(L) EFFYSYMTGL	LRPDS
1	Q9H245	ck: 5880 len: 531 i Q9h245 homo sapiens (human). clorf28. 12 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(L) PVFLRFWEVTL	DRYVY
1	509: RHLDR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(R)(F)WXX(L) PVFLRFWEVTL	DRYVY
1	Q9BXJ8	ck: 1303 len: 343 i Q9bxj8 homo sapiens (human). transmembra (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(F) LFFGFHWQLF	NALTL
1	279: FLLEP	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(F) LFFGFHWQLF	NALTL
1	Q9NV66	ck: 4753 len: 732 i Q9nv66 homo sapiens (human). cdna flj109 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLEKDFWQRF	LDSLK
1	547: KRIDR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLEKDFWQRF	LDSLK
1	Q9H9C4	ck: 6460 len: 561 i Q9h9c4 homo sapiens (human). cdna flj128 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLEKDFWQRF	LDSLK
1	376: KRIDR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(D)(F)WXX(F) PLEKDFWQRF	LDSLK

1	Q9H5L7	ck: 1905	len: 355	1	Q9H517 homo sapiens (human) . cdna: flj23316	275: YSSVF	(L)xFx(S)(F)Wxx(L)	LSFVSFWIDL	KALPA	
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)								
		(P)xFx(R)(F)Wxx(L)								
	333: RHLDL	PVFLRFMETL							DRYVW	
1	Q96KN6	ck: 6589	len: 1,449	1	Q96kn6 homo sapiens (human) . putative trp c	Q9XYS1	ck: 646	len: 438	1	Q9xys1 ciona intestinalis. brachyury pro
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)								
		(L)xFx(D)(F)Wxx(L)								
	842: MKKAA	LYFSDFWNKL								
1	Q96JM4	ck: 5997	len: 1,227	1	Q96jma homo sapiens (human) . klal1801 prote	110: IHPDS	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	PNGGSHMWKQ	PIGFS	
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)								
		(E)xFx(S)(Y)Wxx(L)								
	942: SISTV	EAFFSYWLP				106: MHPDS	(P)xFx(S)(H)Wxx(O)	PNGGSHMWKQ	PLSFS	
1	Q96A07	ck: 1525	len: 355	1	Q96a07 homo sapiens (human) . similar to hpf	Q9NHK6	ck: 7692	len: 110	1	Q9nhk6 plasmodium falciparum. pfemp1 pro
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)								
		(P)xFx(R)(F)Wxx(L)								
	333: RHLDL	PVFLRFMETL				47: DKLSL	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	EKFREYMWAL	NRKEV	
1	Q9V696	ck: 9003	len: 2,040	1	Q9v696 drosophila melanogaster (fruit fly) .	Q9NHG3	ck: 7219	len: 110	1	Q9nhg3 plasmodium falciparum. pfemp1 pro
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)								
		(E)xFx(D)(H)Wxx(O)								
	749: IFLVL	EHFSDHMHKQ				46: KRLLT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	EQFREYMWAL	NRQVQ	
1	Q9V9K3	ck: 7480	len: 337	1	Q9v9k3 drosophila melanogaster (fruit fly) .	Q9U2I8	ck: 2401	len: 569	1	Q9u2i8 caenorhabditis elegans. y41c4a.2
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)								
		(E)xFx(E)(F)Wxx(L)								
	183: YPTAE	EDPDEFWNVL				442: IESYK	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	LKFIEYWKSL	GVDVL	
1	Q9GN87	ck: 9284	len: 209	1	Q9gn87 anopheles dirus, and anopheles dirus	Q9U2X4	ck: 4393	len: 331	1	Q9u2x4 caenorhabditis elegans. y113g7b.9
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)								
		(Q)xFx(D)(Y)Wxx(O)								
	106: MGLTY	QRFQDIWTFQ				167: GCATN	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	QCFPHYWLIO	RSIIF	
1	Q9NL42	ck: 320	len: 491	1	Q9nl42 clona savignyi. brachyury. 12/2001	Q9NB09	ck: 7148	len: 301	1	Q9nbu9 trypanosoma brucei. possible h-sc
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)								
		(P)xFx(S)(H)Wxx(O)								
	109: IHPDS	PNGGSHMWKQ				124: YITDK	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	ELFQDHWTLL	YFGFS	
1	Q9BLF5	ck: 1820	len: 308	1	Q9blf5 olkopleura longicauda. brachyury pro	Q9NB66	ck: 5512	len: 136	1	Q9nb66 trypanosoma brucei. hypothetical1
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)								
		(P)xFx(S)(H)Wxx(O)								
	108: MHPDS	PNGGSHMWKQ				59: KNLFT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	LLFHEWFPL	CLFVS	
1	Q9BLF4	ck: 7868	len: 403	1	Q9blf4 olkopleura longicauda. brachyury pro	Q942I6	ck: 5664	len: 589	1	Q942i6 caenorhabditis elegans. hypothetical1
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)								
		(P)xFx(S)(H)Wxx(O)								
	108: MHPDS	PNGGSHMWKQ				222: TRBDV	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	ENFYDFWFNF	QSWRE	
1	O17041	ck: 8670	len: 491	1	O17041 caenorhabditis elegans. hypothetical1	0462I8	ck: 7893	len: 219	1	0462i8 culicoides varilipennis. glutathio
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)								
		(P)xFx(S)(H)Wxx(O)								
		PNGGSHMWKQ				108: MGLTY	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	QRFADYWFQ	LFQKO	

1	Q17999	ck: 7776	len: 814	1	Q17999	caenorhabditis elegans. c15a7.1 prot	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	LSPDH
1	106: FFRTR	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(D)(F)WXX(Q)	LEFVDFWLYQ	IMKPR			
1	Q18843	ck: 7424	len: 660	1	Q18843	caenorhabditis elegans. c54g10.3 prc	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	LSPDH
1	9: LALDF	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(R)(F)WXX(L)	LEFRFRWRRL	KILFP			
1	Q9XUC9	ck: 6239	len: 170	1	Q9XUC9	caenorhabditis elegans. ml62.2 prote	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	MMPDH
1	83: GSIFG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(E)XFX(H)(F)WXX(L)	EGEKHFWMIGL	NRDPI			
1	Q45282	ck: 676	len: 705	1	Q45282	caenorhabditis elegans. c31a11.5 prc	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	MNDP
1	294: IESIL	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(P)XFX(D)(F)WXX(L)	PFRGFRWNHL	VMNAF			
1	Q01616	ck: 9544	len: 533	1	Q01616	caenorhabditis elegans. hypothetical	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	GIGGD
1	91: LKLET	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(Q)XFX(D)(F)WXX(Q)	QTEPDWDSDQ	LMNPP			
1	Q04622	ck: 3173	len: 326	1	Q04622	trypanosoma brucei. u2 small	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	GIGGD
1	129: RELRD	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(S)(H)WXX(L)	LLEFLSHMKKL	EILSI			
1	Q95JP5	ck: 5683	len: 235	1	Q95JP5	macaca fascicularis (crab eating mac	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	GIGGD
1	16: SCVPI	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(S)(Y)WXX(F)	LVPSPYMASF	SCCMW			
1	Q9TGI4	ck: 134	len: 161	1	Q9TGI4	taenia crassiceps. nadh dehydrogenas	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	GIGGD
1	147: SYEYF	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(C)(F)WXX(L)	LVSFCFWCYL	GFFIF			
1	Q9B8W7	ck: 7748	len: 417	1	Q9B8W7	taenia crassiceps. nadh-ubiquinone c	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	GIGGD
1	403: SYEYF	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(C)(F)WXX(L)	LVSFCFWCYL	GFFIF			
1	Q94TE9	ck: 2807	len: 116	1	Q94TE9	chlorophthalmus agassizii (shortnose	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	GIGGD
1	18: ALSTL	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(S)(F)WXX(O)	LAFVSTWLPQ	LNPDT			
1	Q94T60	ck: 3118	len: 116	1	Q94T60	trachipterus trachipterus (ribbon fi	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	GIGGD

1	18: TLSTV	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(S)(F)WXX(Q)	LIFISFWLPQ	LSPDH		
1	Q94T47	ck: 2171	len: 116	1	Q94T47	zu cristatus (scalloped ribbonfis	
1	18: LLSV	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(S)(F)WXX(Q)	LIFISFWLPQ	LSPDH		
1	Q94S17	ck: 2075	len: 116	1	Q94S17	antigonla capros (deepbody boarl	
1	18: TLSTI	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(S)(F)WXX(Q)	LAFVSTWLPQ	MMPDH		
1	Q94SD5	ck: 3914	len: 116	1	Q94SD5	dactyloptena peterseni (starry fl	
1	18: LLTTV	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(S)(F)WXX(Q)	LAFVSTWLPQ	MNDP		
1	Q96303	ck: 605	len: 534	1	Q96303	arabidopsis thaliana (mouse-ear c	
1	130: AVMAT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(R)(F)WXX(F)	LCFFRFWLG	GIGGD		
1	Q50040	ck: 992	len: 542	1	Q50040	arabidopsis thaliana (mouse-ear c	
1	130: GVMAT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(R)(F)WXX(F)	LCFFRFWLG	GIGGD		
1	Q96302	ck: 8746	len: 524	1	Q96302	arabidopsis thaliana (mouse-ear c	
1	130: GVMTP	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(R)(F)WXX(F)	LCFFRFWLG	GIGGD		
1	Q48640	ck: 8027	len: 524	1	Q48640	arabidopsis thaliana (mouse-ear c	
1	130: GVMTP	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(R)(F)WXX(F)	LCFFRFWLG	GIGGD		
1	Q48639	ck: 4821	len: 521	1	Q48639	arabidopsis thaliana (mouse-ear c	
1	130: GVMTP	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(R)(F)WXX(F)	LCFFRFWLG	GIGGD		
1	Q92GX5	ck: 8548	len: 3,449	1	Q92GX5	arabidopsis thaliana (mouse-ear c	
1	70: KKLKE	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(R)(Y)WXX(F)	LDKRYMEEF	RSSSS		
1	Q92WT3	ck: 9638	len: 516	1	Q92WT3	arabidopsis thaliana (mouse-ear c	
1	131: TVMVT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)	(L)XFX(R)(F)WXX(F)	LCFFRFWLG	GIGGD		

1	Q9SQY1	ck: 9617	len: 558	1	Q9sqy1	arabidopsis thaliana (mouse-ear cress)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
1	535: SAMAM					LSVSR
		(L)XFX(C)(F)WXX(L)				
		LIFMCFWVLL				
1	Q9AU01	ck: 5248	len: 540	1	Q9au01	lupinus albus (white lupine). phosphatase
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	130: SYLAT					GIGGD
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	Q9AU00	ck: 478	len: 539	1	Q9au00	lupinus albus (white lupine). phosphatase
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	130: GTVST					GIGGD
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	Q9AVR0	ck: 6827	len: 536	1	Q9avr0	sesbania rostrata. phosphate transp
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	129: GVMAT					GIGGD
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	Q9AVQ9	ck: 3365	len: 540	1	Q9avq9	sesbania rostrata. phosphate transp
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	132: GVMAT					GIGGD
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	Q96243	ck: 8141	len: 524	1	Q96243	arabidopsis thaliana (mouse-ear cress)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	130: GVMAT					GIGGD
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	Q96264	ck: 8941	len: 524	1	Q96264	arabidopsis thaliana (mouse-ear cress)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	130: GVMAT					GIGGD
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	Q24029	ck: 4193	len: 538	1	Q24029	lycopersicon esculentum (tomato). ph
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	129: GVMAT					GIGGD
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	004660	ck: 5621	len: 938	1	004660	arabidopsis thaliana (mouse-ear cress)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(S)(F)WXX(F)				
		LGPDSEFWL				
1	856: VSEFRQ					LVAAT
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(S)(F)WXX(F)				
		LGPDSEFWL				
1	022055	ck: 2442	len: 542	1	022055	catharanthus roseus (rosy periwinkle)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	130: AVMAT					GIGGD
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	Q9ST22	ck: 2750	len: 536	1	Q9st22	nicotiana tabacum (common tobacco).
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(F)				
		LCFFRFWLG				
1	129: GVMAT					GIGGD
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	Q9FH47	ck: 6231	len: 1,112	1	Q9fh47	arabidopsis thaliana (mouse-ear cress)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(Q)XFX(S)(Y)WXX(L)				
		QNRSTWGL				
1	188: SKDDF					DYKIK
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(Q)XFX(S)(Y)WXX(L)				
		QNRSTWGL				
1	Q9AYT3	ck: 5587	len: 537	1	Q9ayt3	nicotiana tabacum (common tobacco)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	130: SVMTT					GIGGD
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	Q9AYT2	ck: 7879	len: 523	1	Q9ayt2	nicotiana tabacum (common tobacco)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	131: GVMAT					GIGGD
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	Q9AYT1	ck: 9177	len: 523	1	Q9ayt1	nicotiana tabacum (common tobacco)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	131: GVMAT					GIGGD
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(F)				
		LCFFRFWLG				
1	065383	ck: 9544	len: 783	1	065383	arabidopsis thaliana (mouse-ear cress)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(L)				
		LYFGRFWTHL				
1	158: NGSDS					VNAYA
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(F)WXX(L)				
		LYFGRFWTHL				
1	022148	ck: 8740	len: 1,374	1	022148	arabidopsis thaliana (mouse-ear cress)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(H)WXX(L)				
		LGFRDHIRL				
1	596: KAMRG					IMECV
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(H)WXX(L)				
		LGFRDHIRL				
1	082193	ck: 6852	len: 550	1	082193	arabidopsis thaliana (mouse-ear cress)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(Y)WXX(L)				
		LSFGRTWNYL				
1	280: PLASD					RKQTF
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(R)(Y)WXX(L)				
		LSFGRTWNYL				
1	082362	ck: 3197	len: 787	1	082362	arabidopsis thaliana (mouse-ear cress)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(F)				
		LOFMDFWLSF				
1	629: LSYNV					MAVVG
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(L)XFX(D)(F)WXX(F)				
		LOFMDFWLSF				
1	Q9ZQ76	ck: 6617	len: 131	1	Q9zq76	arabidopsis thaliana (mouse-ear cress)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P)XFX(S)(H)WXX(L)				
		PSFRSHWKS				
1	49: FLRKS					KORNL
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P)XFX(D)(F)WXX(F)				
		PSFRSHWKS				
1	Q9FRL1	ck: 7509	len: 654	1	Q9frl1	arabidopsis thaliana (mouse-ear cress)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				
		(P)XFX(D)(F)WXX(F)				
		PSFRSHWKS				

470: KAIDR	PLFADFMEERF	IDSIX			
022301	ck: 33 len: 537 1 022301 medicago truncatula (barrel medic).		1	09M4X4	ck: 7892 len: 709 1 09m4x4 zea mays (maize). transposase. 6/
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(L)XFX(R)(F)WXX(F)			434: FIHRK	(L)XFX(R)(F)WXX(F)
129: SVMAT	LCFFRFWLG	GIGGD			LCFFRFWLG
022302	ck: 2173 len: 533 1 022302 medicago truncatula (barrel medic).		1	09LKL5	ck: 9448 len: 390 1 09lkl5 oryza sativa (rice). phosphate tr
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(L)XFX(R)(F)WXX(F)			99: GVMAT	(L)XFX(R)(F)WXX(F)
129: SVMAT	LCFFRFWLG	GIGGD			LCFFRFWLG
022548	ck: 4793 len: 538 1 022548 lycopersicon esculentum (tomato). ir		1	09ATE1	ck: 2612 len: 323 1 09ate1 oryza sativa (rice). phosphate tr
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(L)XFX(R)(F)WXX(F)			92: SVMAT	(L)XFX(R)(F)WXX(F)
129: GVMAT	LCFFRFWLG	GIGGD			LCFFRFWLG
022549	ck: 3242 len: 528 1 022549 lycopersicon esculentum (tomato). ir		1	09AXE0	ck: 5023 len: 333 1 09axe0 oryza rufipogon (wild rice). phos
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(L)XFX(R)(F)WXX(F)			92: GVMAT	(L)XFX(R)(F)WXX(F)
135: GVMAT	LCFFRFWLG	GIGGD			LCFFRFWLG
09XEL6	ck: 7287 len: 396 1 09xel6 triticum aestivum (wheat). high-affi		1	09M1T0	ck: 3263 len: 535 1 09m1t0 arabidopsis thaliana (mouse-ear c
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(L)XFX(R)(F)WXX(F)			130: TVMTT	(L)XFX(R)(F)WXX(F)
99: SVMAT	LCFFRFWLG	GIGGD			LCFFRFWLG
092PFA	ck: 2196 len: 382 1 092pfa arabidopsis thaliana (mouse-ear cres		1	09LZ39	ck: 3770 len: 428 1 09lzf39 arabidopsis thaliana (mouse-ear c
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(L)XFX(R)(F)WXX(F)			182: FFLCY	(L)XFX(C)(F)WXX(F)
112: TVMTT	LCFFRFWLG	GIGGD			LSFLCFWLG
095WG9	ck: 8707 len: 540 1 095wg9 solanum tuberosum (potato). inorgan		1	09ASF0	ck: 5843 len: 794 1 09asf0 oryza sativa (rice). hypothetical
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(L)XFX(R)(F)WXX(F)			263: KATAL	(L)XFX(R)(F)WXX(Q)
129: SVMAT	LCFFRFWLG	GIGGD			LYFERFMWQ
09ILS5	ck: 6167 len: 537 1 09ils5 nicotiana tabacum (common tobacco).		1	09ARI9	ck: 1485 len: 543 1 09ari9 lupinus albus (white lupine). pho
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(L)XFX(R)(F)WXX(F)			130: SVLAT	(L)XFX(R)(F)WXX(F)
130: SVMAT	LCFFRFWLG	GIGGD			LCFFRFWLG
09M5K0	ck: 8753 len: 265 1 09m5k0 oryza sativa (rice). phosphate trans		1	040115	ck: 3380 len: 214 1 040115 luffa cylindrica (smooth loofah)
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(L)XFX(R)(F)WXX(F)			24: PMVKS	(Q)XFX(S)(F)WXX(Q)
99: GVMAT	LCFFRFWLG	GIGGD			QTFDSFMWQ
09M562	ck: 15 len: 393 1 09m562 oryza sativa (rice). phosphate trans		1	004381	ck: 4865 len: 521 1 004381 arabidopsis thaliana (mouse-ear c
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)				(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(L)XFX(R)(F)WXX(F)			130: GVMAT	(L)XFX(R)(F)WXX(F)
99: GVMAT	LCFFRFWLG	GIGGD			LCFFRFWLG
			1	043650	ck: 8926 len: 540 1 043650 solanum tuberosum (potato). inorg
					(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)

129: SVMGT	(L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD	129: SVMGT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD
Q41479 ck: 2263 len: 527	1 Q41479 solanum tuberosum (potato). inorganic	1	130: GVMGT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD
082003 ck: 5788 len: 538	1 082003 lycopersicon esculentum (tomato). pu	1	129: GVMGT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD
Q94DB8 ck: 3975 len: 555	1 Q94db8 oryza sativa (rice). putative phosph	1	136: AVIGT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD
Q949N0 ck: 1025 len: 549	1 Q949n0 arabidopsis thaliana (mouse-ear cres	1	526: SAMAM	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(L) LTFMCFWVL	LSVSF
Q946Y1 ck: 9530 len: 520	1 Q946y1 gossypium arboreum. ceramide glucosy	1	306: PLASD	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(L) LSFSRYWNL	RKQTF
Q945E6 ck: 7000 len: 538	1 Q945e6 oryza sativa subsp. indica. phosphat	1	129: GVMAT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD
Q944R4 ck: 3251 len: 519	1 Q944r4 arabidopsis thaliana (mouse-ear cres	1	305: PLASD	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(L) LSFGRYWNL	RKQTF
Q943Y7 ck: 3376 len: 125	1 Q943y7 thnopyrum intermedium. inorganic ph	1	93: GVMGT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GVGCD
Q93WR0 ck: 9360 len: 525	1 Q93wr0 triticum aestivum (wheat). inorganic	1	130: GVMGT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GVGCD
Q93WQ9 ck: 9330 len: 525	1 Q93wq9 triticum aestivum (wheat). inorganic	1	130: GVMGT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GVGCD
Q93M08 ck: 8781 len: 525	1 Q93m08 triticum aestivum (wheat). inorga	1	130: GVMGT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GVGCD
Q9D6E1 ck: 164 len: 179	1 Q9d6e1 mus musculus (mouse). 2900075g08r	1	116: GLNVP	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(F) LTFYHFWRYF	HCPAD
Q9C2I3 ck: 6690 len: 139	1 Q9c2i3 mus musculus (mouse). 2310040a02r	1	13: LSVAG	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (C)XFX(R)(H)WXX(L) CNFLRHWLVL	STLSC
Q63363 ck: 7377 len: 475	1 Q63363 rattus norvegicus (rat). rat noer	1	76: IQORP	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(E)(Y)WXX(L) EQFVEHMKEL	KQLMA
Q9DCJ5 ck: 5548 len: 172	1 Q9dcj5 mus musculus (mouse). 0610033103r	1	80: KSHCA	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(E)(Y)WXX(L) EPFETWTC	DYSNM
Q91YD4 ck: 1512 len: 1,507	1 Q91yd4 mus musculus (mouse). transient r	1	860: MKMAS	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(F)WXX(L) LTFSDWNL	DVGAI
056074 ck: 8233 len: 3,005	1 056074 hepatitis gb virus a. genome poly	1	789: GRLRK	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(S)(Y)WXX(F) LEFYSYWCRP	HOAYR
Q9DM56 ck: 4069 len: 655	1 Q9dm56 rat cytomegalovirus (strain maast	1	210: GRALT	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(E)(H)WXX(L) LRPEHMYLL	POPMN
Q83896 ck: 5502 len: 595	1 Q83896 ovine adenovirus type 7. terminal	1	241: ECILD	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(E)(F)WXX(L) LPSEFWIEL	FIDFR
Q91GE6 ck: 8973 len: 243	1 Q91ge6 epiphyas postvittana nucleopolyhe	1	51: TSFVE	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(E)(Y)WXX(L) LOFLEWYML	SLLIN
Q9W7A2 ck: 3716 len: 380	1 Q9w7a2 coturnix japonica (japan	1			

1		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (C)XFX(S)(F)WXX(L) CCFSSFWKGL	VASSV		
319:	CVGTC				
1	Q90ZD4	ck: 3654 len: 491	1	Q90zd4	gallus gallus (chicken). interferon
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(E)(Y)WXX(Q) PQCEYWEQ	VORGG		
446:	LVKLV				
1	Q25915	ck: 8156 len: 417	1	Q25915	helicobacter pylori (campylobacter f
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(S)(H)WXX(F) QDFKSHWRHF	IDSAH		
108:	SLTRK				
1	P72685	ck: 1310 len: 290	1	P72685	synechocystis sp. (strain pcc 6803).
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(S)(H)WXX(Q) LIFFSHMCYQ	TPSL		
68:	IFLVL				
1	Q9PAP6	ck: 144 len: 544	1	Q9PAP6	xylella fastidiosa. hypothetical pro
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(S)(H)WXX(F) LRFSSHWFVF	LLSVA		
14:	VDSMN				
1	Q9KVA4	ck: 1939 len: 463	1	Q9KVA4	vibrio cholerae. rfbg protein. 12/20
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(F)WXX(F) EKFSFMDPF	VFEYN		
161:	HLGG				
1	Q9KQ57	ck: 2237 len: 118	1	Q9KQ57	vibrio cholerae. hypothetical protei
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(S)(Y)WXX(Q) LGFTSLMDQ	NS		
107:	LALSM				
1	Q9KE45	ck: 7616 len: 155	1	Q9KE45	bacillus halodurans. bh1013 protein.
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(H)WXX(L) ERFDDHWVL	SANHG		
43:	CAMTA				
1	Q9KB88	ck: 9494 len: 818	1	Q9KB88	bacillus halodurans. bh2040 protein.
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(S)(Y)WXX(Q) QLFMSYMAVQ	SLANY		
235:	GPIKR				
1	Q9K650	ck: 5308 len: 154	1	Q9K650	bacillus halodurans. putative rna po
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(L) LVFKHFWLGL	TSQEI		
110:	ASLSD				
1	Q9JY55	ck: 3936 len: 167	1	Q9JY55	neisseria meningitidis (serogroup b)
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(C)(Y)WXX(L) PYFACYWGL	CYPVS		
140:	EGGMM				
1	Q9JVL8	ck: 1872 len: 173	1	Q9JVL8	neisseria meningitidis (serogroup
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(C)(Y)WXX(L) PFICYWGL	FYPAN		
146:	ERGMM				
1	Q9JY37	ck: 492 len: 172	1	Q9JY37	neisseria meningitidis (serogroup
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (P)XFX(C)(Y)WXX(L) PYFACYWGL	CYPVS		
145:	EGGMM				
1	P72121	ck: 9154 len: 723	1	P72121	pseudomonas aeruginosa. outer mem
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(Y)WXX(F) BRFPDYWELF	SPRKG		
491:	GLGHA				
1	Q916K4	ck: 1486 len: 392	1	Q916K4	pseudomonas aeruginosa. probable
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(R)(F)WXX(L) QHFFRFWLML	TTGMM		
49:	LHPAL				
1	Q915J3	ck: 7512 len: 358	1	Q915J3	pseudomonas aeruginosa. hypotheti
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(S)(F)WXX(F) LLFYRFGLF	PTVAQ		
89:	AASYF				
1	Q9CNB6	ck: 8093 len: 1,230	1	Q9CNB6	pasteurella multocida. recb. 12/2
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(H)WXX(L) ETFSHNMCL	QYVLT		
856:	VNALP				
1	Q9AAB3	ck: 7467 len: 208	1	Q9AAB3	caulobacter crescentus. hypotheti
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(F)WXX(L) EGFEDFWGL	QGETD		
58:	VNGCG				
1	Q98JY9	ck: 592 len: 153	1	Q98JY9	rhizobium lot1 (mesorhizobium lot
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(L) LMFSRYWML	ASRAA		
113:	HCSSR				
1	Q97KS8	ck: 8409 len: 260	1	Q97KS8	clostridium acetobutylicum. abc-t
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(H)WXX(F) EKFSHMGHF	SLTWY		
33:	FSFNS				
1	Q97DB6	ck: 4438 len: 577	1	Q97DB6	clostridium acetobutylicum. misma
		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(R)(Y)WXX(F) ERFRYWGEP	QKRKR		
64:	HKLNY				
1	Q99VP6	ck: 6701 len: 501	1	Q99VP6	staphylococcus aureus (strain n31


```

1      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
29: RGLGV      (L)XFX(E)(F)WXX(F)      STYGM
      LFFVEFWERE

1      Q92AV1  ck: 1480  len: 289   1 Q92av1 listeria innocua. 11n1818 protein. 1
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
175: HLNFS      (E)XFX(S)(F)WXX(L)      RMTVT
      EDFLSFWNML

1      Q929W8  ck: 3736  len: 391   1 Q929w8 listeria innocua. 11n2155 protein. 1
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
240: YRGPF      (L)XFX(D)(F)WXX(L)      KFRLI
      LSFEDYWPPL

1      Q928I2  ck: 2291  len: 510   1 Q928i2 listeria innocua. pgm protein. 12/20
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
31: AQANK      (P)XFX(R)(Y)WXX(F)      PHGEL
      PNFDRYMANF

1      Q29337  ck: 7306  len: 315   1 Q29337 archaeoglobus fulgidus. hypothetical
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
53: AVYIL      (L)XFX(S)(F)WXX(L)      SNLEF
      LFFWSFWNLL

1      Q29916  ck: 4349  len: 396   1 Q29916 archaeoglobus fulgidus. pheromone sh
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
116: DRDIG      (L)XFX(R)(F)WXX(L)      TFEVK
      LFTFRFWOKL

1      Q9YFF0  ck: 4613  len: 110   1 Q9yff0 aeropyrum pernix. 110aa long hypothe
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
96: SDVEG      (E)XFX(S)(F)WXX(Q)      RPRTD
      EAFKSEWGAQ

1      Q9HIS0  ck: 3314  len: 150   1 Q9his0 thermoplasma acidophilum. hypothetical
      (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
26: LGLRR      (L)XFX(R)(F)WXX(F)      WLDDH
      LDFREWLME

```

Databases searched:

SWISS-PROT, Release 40.1, Released on 11Feb2002, Formatted on 19Feb2002
 SPTREMBL, Release 19.0, Released on 14Dec2001, Formatted on 19Feb2002

Total finds: 174
 Total length: 211,714,479
 Total sequences: 667,446
 CPU time: 23:58.58

This Page Blank (uspio)

1 FINDPATTERNS on pir:* allowing 0 mismatches			
1	1 (P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		May
1	G71532 ck: 4900 len: 934 1 probable threonine/tyrosine-specific prote	1	T32253 ck: 8670 len: 491 1 hypothetical protein T1587.16 - Caenorha
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(E)XFX(S)(F)WXX(F)		(L)XFX(S)(F)WXX(L)
1	664: RSSKM		275: YSSVF
	ELFLSFWTGF		LSFVSFWIDL
	SLFLP		KALPA
1	W5WLRI ck: 8739 len: 157 1 E5 protein - rhesus papillomavirus (type 1)	1	S40462 ck: 4989 len: 535 1 t-complex-type molecular chaperone tcpl
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(P)XFX(C)(F)WXX(F)		(L)XFX(E)(Y)WXX(F)
	PPCCFWLGF		LDPEYWRPF
1	37: YACKP		3: MA
	CCCCC		ILIRE
1	C69762 ck: 9486 len: 492 1 dl-tripeptide ABC transporter (membrane pr	1	PC2306 ck: 7087 len: 189 1 dnaJ protein - Synechococcus sp. (strain
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(L)XFX(E)(F)WXX(F)		(O)XFX(R)(Y)WXX(O)
	LEFTFWEF		QRFSTRWQQ
1	28: RGLFT		68: RRAEY
	SYGGM		GAASV
1	C69291 ck: 4349 len: 396 1 pheromone shutdown protein (crab) homolog	1	JC5550 ck: 2256 len: 287 1 dnaJ protein - Synechococcus sp.
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(L)XFX(R)(F)WXX(L)		(O)XFX(R)(Y)WXX(O)
	LTFTRFWQKL		QRFSTRWQQ
1	116: DRDIG		68: RRAEY
	TFVEK		GAASV
1	G86502 ck: 36 len: 918 1 S/T protein kinase [imported] - Chlamydophi	1	JC5648 ck: 4120 len: 641 1 terminal protein precursor - Ovline adeno
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(E)XFX(S)(Y)WXX(F)		(L)XFX(E)(F)WXX(L)
	ELFLSYWSGF		LPPSEWIEL
1	648: RSSKM		287: BCILD
	IPHLN		FIDBF
1	C72120 ck: 36 len: 918 1 s/t protein kinase - Chlamydoiphia pneumoni	1	T08841 ck: 8233 len: 3,005 1 polypeptin - douroucouli hepatitis GB v
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(E)XFX(S)(Y)WXX(F)		(L)XFX(S)(Y)WXX(F)
	ELFLSYWSGF		LEFYSYCRF
1	648: RSSKM		789: GRLRK
	IPHLN		HQAVR
1	AF1811 ck: 1483 len: 356 1 dTDP-glucose 4-6-dehydratase [imported] - A	1	H89844 ck: 6701 len: 501 1 hypothetical protein SA0682 [imported] -
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(E)XFX(H)(H)WXX(O)		(L)XFX(E)(F)WXX(F)
	EAFRHHWTKQ		LEFVEFWERF
1	125: TFTLL		29: RGLGV
	GKPRN		SYGGM
1	AH1381 ck: 2157 len: 510 1 phosphoglycerate mutase homolog pgm [import	1	C97003 ck: 8409 len: 260 1 ABC-type spermidine/putrescine transport
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(P)XFX(R)(Y)WXX(F)		(E)XFX(S)(H)WXX(F)
	PNFDRIMADF		EKFSSIMGHF
1	31: AQANK		33: FSEFNS
	PHGEL		SLTWY
1	A11750 ck: 2291 len: 510 1 phosphoglycerate mutase homolog pgm [import	1	C84811 ck: 605 len: 534 1 phosphate transporter (AtPT2) [imported]
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(P)XFX(R)(Y)WXX(F)		(L)XFX(R)(F)WXX(F)
	PNFDRIMADF		LCFFRWLGF
1	31: AQANK		130: AVMAT
	PHGEL		GIGGD
1	A27671 ck: 7377 len: 475 1 spectrin alpha chain, nonerythroid - rat (f	1	T07607 ck: 2263 len: 527 1 phosphate transport protein 2 - potato
	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)
	(E)XFX(E)(H)WXX(L)		(L)XFX(R)(F)WXX(F)
	EOFVEHWKEL		LCFFRWLGF
76: IQORP	KOLAA	133: VYMTT	GIGGD
		T07604 ck: 8926 len: 540 1 phosphate transport protein PT1 - potato	

1		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD	91: LKLETF	QGFPEFWDSDQ	LMNPP
129: SVMAT				556753	ck: 3644 len: 491	i interferon regulatory factor 3 - chicken
1	107740	ck: 4193 len: 538	i probable inorganic phosphate transport prot	1	446: LVKLV	
129: GVMAT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD	573432	ck: 4533 len: 518	i MG096 homolog D09_orf518 - Mycoplasma pn
1	101124	ck: 992 len: 542	i probable phosphate transporter At2g32830 [1	1	251: SLMNF	
130: GVMAT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD	573923	ck: 7095 len: 417	i MG288 homolog G07_orf417 - Mycoplasma pn
1	107892	ck: 33 len: 537	i probable inorganic phosphate transport prot	1	164: SLTSA	
129: SVMAT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD	B72719	ck: 4613 len: 110	i probable bacterioferritin comigratory pr
1	107894	ck: 2173 len: 533	i probable inorganic phosphate transport prot	1	96: SDVEG	
129: SVMAT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD	G70077	ck: 7683 len: 153	i hypothetical protein yx1k - Bacillus sub
1	105714	ck: 4793 len: 538	i probable inorganic phosphate transport prot	1	32: ELEDA	
129: GVMAT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD	T29895	ck: 5664 len: 589	i hypothetical protein F38a5.13 - Caenorha
1	107164	ck: 5788 len: 538	i probable inorganic phosphate transport prot	1	222: TPREDV	
129: GVMAT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD	T19595	ck: 676 len: 705	i hypothetical protein C31A11.5 - Caenorha
1	105724	ck: 3242 len: 528	i probable inorganic phosphate transport prot	1	294: IESIL	
135: GVMAT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD	B81828	ck: 492 len: 172	i hypothetical integral membrane protein N
1	107808	ck: 2442 len: 542	i probable inorganic phosphate transport prot	1	145: EGGWM	
130: AVMAT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD	D81049	ck: 3936 len: 167	i hypothetical protein NM1733 [imported]
1	147629	ck: 3263 len: 535	i phosphate transport protein - Arabidopsis t	1	140: EGGWM	
130: TVMTT		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) LCFFRFWLGF	GIGGD	A81923	ck: 1872 len: 173	i probable membrane protein NMA0784 [impor
1	T34457	ck: 9544 len: 533	i hypothetical protein T19H12.10 - Caenorhabd	1	146: ERGWM	
1		(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F) (Q)XFX(D)(F)WXX(O)				

1	H83554	ck: 7512	len: 358	1	hypothetical protein PA0736 [imported] - <i>Pa</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(F)WXX(F)	PTVAQ
1	89: AASVF					LFYRFWGLF	
1	F83173	ck: 9154	len: 723	1	outer membrane protein OprC PA3790 [imported] - <i>Pa</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(Y)WXX(F) ERPDYWELE	SPKRG
1	491: GLGHA						
1	F83610	ck: 1486	len: 392	1	probable fatty acid desaturase PA0286 [imported] - <i>Pa</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(R)(F)WXX(L) QHFRFWLWL	TTGMN
1	49: LHPAL						
1	S73657	ck: 3241	len: 260	1	MG288 homolog F04_orf260V - <i>Mycoplasma pneumoniae</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(E)(Y)WXX(Q) QAEKEYWQKQ	KKDVT
1	2: M						
1	E69365	ck: 7306	len: 315	1	hypothetical protein AF0925 - <i>Archaeoglobus fulgidus</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(S)(F)WXX(L) LFFWSEFWLL	SNLTF
1	53: AVYIL						
1	T01809	ck: 5621	len: 938	1	hypothetical protein A_TM021B04.3 - <i>Arabidopsis thaliana</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(S)(F)WXX(F) LGFDSFWLFE	LVAAL
1	856: VSFRO						
1	T01083	ck: 8548	len: 3,449	1	hypothetical protein T10P11.5 - <i>Arabidopsis thaliana</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(R)(Y)WXX(F) LDRKRYWEFE	RSSSS
1	70: KILKE						
1	T48284	ck: 3770	len: 428	1	hypothetical protein T22P11.220 - <i>Arabidopsis thaliana</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(C)(F)WXX(F) LSFLCFWLCP	CWNHK
1	182: FFLCY						
1	T39938	ck: 5030	len: 1,649	1	hypothetical protein SPBC23E6.04c - <i>flissionospora</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(Y)WXX(L) EFVSVYWSL	IKSRK
1	337: EHRLD						
1	S36335	ck: 3173	len: 326	1	U2 snRNP 40K protein - <i>Trypanosoma brucei</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(S)(H)WXX(L) LFLSHMKKL	EILSI
1	129: RELRD						
1	T19297	ck: 7776	len: 814	1	hypothetical protein C15A7.1 - <i>Caenorhabditis elegans</i>		
1	106: FTRTR						
1	T23772	ck: 6239	len: 170	1	hypothetical protein M162.2 - <i>Caenorhabditis elegans</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(D)(F)WXX(Q) LEFVDFWLQ	IMKPR
1	83: GSIFG						
1	T26443	ck: 4393	len: 331	1	hypothetical protein Y11367B.9 - <i>Caenorhabditis elegans</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(H)(Y)WXX(Q) QCFHFWLQ	RSIIF
1	167: GCATN						
1	H83904	ck: 9494	len: 818	1	hypothetical protein BH2040 [imported] - <i>Caenorhabditis elegans</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(S)(Y)WXX(Q) QLEMSYWAQ	SLANY
1	235: GPIKR						
1	B84135	ck: 5308	len: 154	1	RNA polymerase ECF-type sigma factor BH3 - <i>Caenorhabditis elegans</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (L)XFX(H)(F)WXX(L) LVFKHEWLGL	TSQET
1	D97337	ck: 4438	len: 577	1	mismatch repair protein Muts-1-like ATPase - <i>Caenorhabditis elegans</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(R)(Y)WXX(F) ERFKRYWGCF	OKRRK
1	64: HRLNY						
1	A10020	ck: 7936	len: 686	1	probable membrane receptor protein YP001 - <i>Caenorhabditis elegans</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(D)(Y)WXX(F) ERFPDYWELE	SPKRG
1	455: GLGHT						
1	AB1288	ck: 1598	len: 289	1	transport proteins homolog lmo1706 [imported] - <i>Caenorhabditis elegans</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(F)WXX(L) EDFLSEFWNL	RWTYT
1	175: HLNFS						
1	A11659	ck: 1480	len: 289	1	transport protein homolog lln1818 [imported] - <i>Caenorhabditis elegans</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (E)XFX(S)(F)WXX(L) EDFLSEFWNL	RWTYT
1	175: HLNFS						
1	AC2420	ck: 9610	len: 313	1	hypothetical protein alr4915 [imported] - <i>Caenorhabditis elegans</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(R)(Y)WXX(L) QNETRYWSIL	GSSSY
1	57: SEELI						
1	AF2463	ck: 4158	len: 59	1	CAB/ELIP/HLP superfamily of protein [imported] - <i>Caenorhabditis elegans</i>	(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L) (Q)XFX(H)(F)WXX(L) OGFLHFWGIL	
1	50: ELFSG						

AG2271 ck: 4679 len: 59 1 CAB/ELIP/HLIP superfamily [imported] - Anah

(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)

50: EAFSG (Q)XFX(H)(F)WXX(L) QGFLHFWGIL

AH1862 ck: 580 len: 40 1 CAB/ELIP/HLIP-related protein [imported] -

(P,L,E,C,Q)XFX(R,H,E,C,S,D)(H,F,Y)WXX(F,Q,L)

31: EEFSG (Q)XFX(H)(F)WXX(L) QGFLHFWNITL

Databases searched: NBRF, Release 71.0, Released on 15Feb2002, Formatted on 19Feb2002

Total finds: 85

Total length: 96,089,334

Total sequences: 283,138

CPU time: 10:43.16

This Page Blank (uspio)

11AA_SEQUENCE 1.0 STANDARD; PRT; 387 AA.
 ID 2A5E_MOUSE
 AC 061151;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Serine/threonine protein phosphatase 2A, 56 kDa regulatory subunit,
 DE epsilon isoform (PP2A, B subunit, B' epsilon isoform) (PP2A, B
 DE subunit, B56 epsilon isoform) (PP2A, B subunit, PR61 epsilon isoform)
 DE (PP2A, B subunit, R5 epsilon isoform) (Fragment).
 GN PPP2R5E.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Embryonic fibroblast;
 RX MEDLINE=97042488; PubMed=8887688;
 RA Okamoto K., Kamabayashi C., Serrano M., Prives C., Mumby M.C.,
 RA Beach D.;
 RT "p53-dependent association between cyclin G and the B' subunit of
 RT protein phosphatase 2A.";
 RL Mol. Cell. Biol. 16:6593-6602(1996).
 RN [2]
 RP IDENTIFICATION OF PROBABLE FRAMESHIFT.
 RA Hulo C.;
 RT Unpublished observations (JAN-2000).
 CC -1- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
 CC SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
 CC LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
 CC COMPARTMENT. INTERACTS WITH CYCLIN G (IN VITRO).
 CC -1- SUBUNIT: PP2A CONSISTS OF A COMMON HETEROODIMERIC CORE ENZYME,
 CC COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa
 CC CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A). THAT ASSOCIATE
 CC WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE
 CC WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
 CC B (THE R2/R3/R5/R55, R3/B'/PR2/PR130/PP59 AND R5/B'/B56
 CC FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS,
 CC AND CELL SIGNALING MOLECULES.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B
 CC FAMILY.
 CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT WAS
 CC INTRODUCED FROM POSITION 112-135 TO EXTEND THE SIMILARITY WITH THE
 CC HUMAN HOMOLOG.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: U49728; AAB37234.1; ALT_FRAME.
 DR MGD: MGI:1349473; Pp2r5e.
 DR InterPro: IPR002554; B56.
 DR Pfam: PF01603; B56; 1.
 DR Multigene family.
 FT NON_TER 1 1
 FT TER 387 387
 SQ SEQUENCE 387 AA; 45497 MW; 780D5404848A548E CRC64;
 2A5E_MOUSE Length: 387 May 30, 2002 09:32 type: P Check: 8649 ..

11AA_SEQUENCE 1.0 STANDARD; PRT; 471 AA.
 ID BRAC_HALRO
 AC P56158;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Brachyury protein homolog (T protein) (AS-T).
 GN T.
 OS Halocynthia roretzi (Sea squirt).
 OC Eukaryota; Metazoa; Chordata; urochordata; Ascidiacea;
 OC Stolidobranchia; Pyridae; Halocynthia.
 OX NCBI_TaxID=7729;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Gastrula;
 RA Yasuo H., Satoh N.;
 RT "An ascidian homolog of the mouse Brachyury (T) gene is expressed
 RT exclusively in notochord cells at the fate restricted stage.";
 RL Dev. Growth Differ. 36:9-18(1995).
 CC -1- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF GENES
 CC REQUIRED FOR MESODERM DIFFERENTIATION (BY SIMILARITY).
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
 CC -1- TISSUE SPECIFICITY: DEVELOPING NOTOCHORD.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED IN 64-CELL EMBRYOS. REACHES
 CC MAXIMAL LEVELS AT THE GASTRULA STAGE, DECREASES IN NEURULA STAGE
 CC AND BARELY DETECTABLE IN LARVAE.
 CC -1- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: D16441; BAA03910.1; -
 DR HSP; P24781; 1YBR.
 DR InterPro: IPR001699; T-box.
 DR Pfam: PF00907; T-box; 1.
 DR PRINTS: PRO0937; TBOX.
 DR SMART: SM00425; TBOX; 1.
 DR PROSITE: PS01283; TBOX_1; 1.
 DR PROSITE: PS01264; TBOX_2; 1.
 DR PROSITE: PS0252; TBOX_3; 1.
 DR Developmental protein; Transcription regulation; DNA-binding;
 KW Nuclear protein; Activator.
 FT DNA_BIND 24 196
 FT DOMAIN 340 343
 FT DOMAIN 348 353
 FT POLY-ASN.
 FT POLY-ASN.
 SQ SEQUENCE 471 AA; 52417 MW; 096BBBD42F205C4 CRC64;
 BRAC_HALRO Length: 471 May 30, 2002 09:32 type: P Check: 462 ..

1 LKDVPTSEOP ELFLKIKQC CVICDMDTL SCLKKREYR STINELVDYI
 51 TISRCLTEQ TYPEVVRMS CNIFRTLPPS DSNEDPREED EPTLEASMPH
 101 LQVYEFIR FWSQEFQPS IAKKIIDOKF VLQLELFDS EDPREYLYK
 151 TVLHRIYGRF LGLRAFIKQ INNIFLRFVY ETEHNGVAE LLEILGSIIN

1 MSTNNMESP SDESVRLTLN DRALWTKFCS LTNEITVKS GRNRPVLYKL
 51 TASGLEPNSM YSFLDFAPA DSNRMKYVNG EWPVGKREP HAASCVYVHP
 101 DSPNGSHMM KQVPSFKVK LFNKNGGQO QIMLSLKY EPRHHVYKV
 151 GEAASERTIA TSPSPESQFI AVTAIONEBY TSLAKKNPF AKAFLDKER
 201 PDQDFHSLA GIPVSPQVP SWYGRNGSTS SARHPTHONS YGGESELTYS

```

251 QDAIPBYTS RNCMRNRYG NARATPYIP HKELTCOATS FEPVPNDGF
301 YPMFNSSEL PRTLENNYSP AMGAYTNSSI VTSSDIOSGN NNNFYSNNN
351 NINTDEVPY TYWTDENSEF YNOSNSGMP GTTYLPYQSS PVNQFSTYCP
401 PYSEIDIS PVOODIINMQ NPYQTAMPP LSYDGCSTMY NSTTPYSSSG
451 ESTTSEMTLL ATARYLQNL R L

!!A SEQUENCE 1.0
ID CHSL_TUBUN STANDARD; PRT; 189 AA.
AC P50003;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Chitin synthase 1 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl
transferase 1) (Fragment).
GN CHSL.
OS Tuber uncinatum (Burgundy truffle).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Pezizomycetes;
OC Pezizales; Tuberaceae; Tuber.
OX NCBI_TaxID=36049;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95031022; PubMed=7944356;
RA Nehman B., Brunner I., Braus G.H.;
RT "Nucleotide sequence variation of chitin synthase genes among
RT ectomycorrhizal fungi and its potential use in taxonomy.";
RT Appl. Environ. Microbiol. 60:3105-3111(1994).
CC -1- FUNCTION: PLAYS A MAJOR ROLE IN CELL WALL BIOGENESIS.
CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + ((1,4)-(N-acetyl-
CC beta-D-glucosaminyl))((N+1).
CC -1- SUBCELLULAR LOCATION: Plasma membrane-bound.
CC -1- SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
-----
CC EMBL: X78101; CNA54994.1; -
CC DR InterPro: IPR002923; Chitin_synth.
CC Pfam: PF01644; Chitin_synth. 1.
CC DR ProDom: PD002998; Chitin_synth. 1.
CC KM Transferase; Glycosyltransferase; Transmembrane; Cell wall;
CC MultiGene family.
CC FT NON_TER 1
CC FT NON_TER 189
CC FT NON_TER 189
CC SO SEQUENCE 189 AA; 21187 MW; 940F6C0ABA7422C9 CRC64;

CHSL_TUBUN Length: 189 May 30, 2002 09:32 Type: P Check: 8541 ..
1 EMMFTRTMG VARNIAHFC SRSRWGKD GWOKIYVCIY ADGRKKVHR
51 VLDAPMGV YQAGIAKNSI NGREYKAHY EYTVQSLDS DLKFKAEKG
101 IYVQMTFL KELNAKKLS HRCIQRFWS TLNPNVCILL DVGTRPGHDS
151 LYHLMKAFDT DSNVGAGCE IKAMKGGM GLENPILVAS

!!A SEQUENCE 1.0
ID DNAJ_SYNP7 STANDARD; PRT; 287 AA.
AC P50026;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chaparone protein dnaJ.

```

```

GN DNAJ.
OS Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=1140;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97382457; PubMed=9240461;
RA Oguchi K., Nimura K., Yoshikawa H., Takahashi H.;
RT "Sequence and analysis of a dnaJ homologue gene in cyanobacterium
RT Synechococcus sp. PCC7942.";
RL Biochem. Biophys. Res. Commun. 236:461-466(1997).
RN [2]
RP SEQUENCE OF 1-189 FROM N.A.
RX MEDLINE=94271242; PubMed=8003021;
RA Nimura K., Yoshikawa H., Takahashi H.;
RT "Sequence analysis of the third dnaK homolog gene in Synechococcus
RT sp. PCC7942.";
RL Biochem. Biophys. Res. Commun. 201:848-854(1994).
RN [3]
RP ERRATUM.
RX MEDLINE=95110355; PubMed=7811295;
RA Nimura K., Yoshikawa H., Takahashi H.;
RL Biochem. Biophys. Res. Commun. 205:2016-2017(1994).
CC -1- FUNCTION: ACTS AS A CO-CHAPERONE. STIMULATES, JOINTLY WITH GRPE,
CC THE ATPASE ACTIVITY OF DNAK (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DNAJ FAMILY.
CC -1- SIMILARITY: CONTAINS 1 J DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
-----
CC EMBL: AB003519; BAA21679.1; -
CC DR EMBL: D29968; BAA06235.1; -
CC DR HSSP: P25685; 1HDJ.
CC DR InterPro: IPR002939; DnaJ_C.
CC DR InterPro: IPR001623; DnaJ_N.
CC DR Pfam: PF00226; DnaJ_1.
CC DR Pfam: PF01556; DnaJ_C_1.
CC DR SMART: SM00271; DnaJ_1.
CC DR PROSITE: PS00636; DnaJ_1; 1.
CC DR PROSITE: PS50076; DnaJ_2; 1.
CC KM Chaparone; DNA replication; Heat shock.
CC FT DOMAIN 4
CC FT DOMAIN 73
CC SO SEQUENCE 287 AA; 31980 MW; 2927A269B2B6C546 CRC64;

DNAJ_SYNP7 Length: 287 May 30, 2002 09:32 Type: P Check: 2256 ..
1 MGNFRDYAL LGIPQADQA AIKAFKRLA RQCHPDLNPG DWAQERFRQ
51 ISAEYELSD PDRAEYORF SRYWQOGAA SVGSDDDYGD FPDFLDFVE
101 LIGRRYERS PRSARSNA TSSALSRDLE RSEYVDPKTA LOGGSAQLDL
151 EDGRLEVDI PAGIQAGEYL RLRGGITKG DLLRLVQLA SNFOYQGSYV
201 IYTLNVSAM AVIGGQYVP TLDPGVOMKL PASIRSGQRL RLAGGYKRP
251 SGRDGDQIV IQLQLPTRL S PEERQLYEQL RSLQSR

!!A SEQUENCE 1.0
ID IRE3_CHICK STANDARD; PRT; 491 AA.
AC Q90643;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Interferon regulatory factor 3 (IRF-3).
OS Gallus gallus (Chicken).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=95334365; PubMed=7541908;
 RA Grant C.E., Vasa M.Z., Deeley R.G.;
 RT "GIRF-3, a new member of the interferon regulatory factor (IRF)
 RL family that is rapidly and transiently induced by dsRNA";
 CC Nucleic Acids Res. 23:2137-2146(1995).
 CC -1- FUNCTION: MAY ACTIVATE TRANSCRIPTION BY COMPLEX FORMATION WITH
 CC OTHER TRANSCRIPTIONAL FACTORS, POSSIBLY MEMBERS OF THE STAT
 CC FAMILY. BINDS SPECIFICALLY TO THE IFN-STIMULATED RESPONSE ELEMENT
 CC (ISRE) BUT NOT TO THE IRF-1 BINDING SITE PRD-1.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U20338; AAA86995.1; -
 DR HSSP: P15314; IIR1
 DR InterPro: IPR001346; IRF.
 DR Pfam: PF00605; IRF; 1.
 DR PRINTS: PR00267; INTERNEGET.
 DR PRODOM: PD002355; IRF; 1.
 DR SMART: SM00348; IRF; 1.
 DR PROSITE: PS00601; IRF; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein; Activator.
 FT DNA_BIND 14 112 TRYPTOPHAN PENTAD REPEAT.
 FT SEQUENCE 491 AA; 54441 MW; CAE0C2AABBE976D9 CRC64;
 SQ
 IRF3_CHICK Length: 491 May 30, 2002 09:32 Type: P Check: 3644 ..

1 MAALSEGDA QKLRGPPYL NAVSSGIYRG LCMIDPRRI FRIPKHNAR
 51 KDVTSSDEVI EKAMAKASGR YEGNAEDPAK WKTNFCALR STMHMLELD
 101 RSKNDNDPRK VYVAASGVN DRSGGPGVAG ALQOQPOLLL NHHDLALENT
 151 PFDSTEGVAA AALTQVDIDL LOSYLOHCNT SALGSOPTLM AHTEGALPED
 201 ALLRGGODGC LPGRFOFDMR QLEEPILIGN OPLTGGGCGQ DGAGALPYSE
 251 ECAIPAPSPA EELFQSANP APPRAGDIG GLPPLIDITI YRGMVYQOE
 301 QVDSRGCVIA YQPLDPAVAE QRLVLFPSPA SLDDPRQRRY TEDLELVAGL
 351 RLEQRAQGLL ATRLKCKVF WALSQOLEGG EPPLNLHRD QETITFDYV
 401 PCTELRDFND SRBRSPDFT IFLCFGQCS STPKRESKLI LKLVLPQCE
 451 YREOVQKRG ASSLSNGVNS LQLSDSFNLF ELITEQYHQT D

111A_SEQUENCE 1.0
 ID PKN1_CHIMU STANDARD: PRT; 934 AA.
 AC 09P692;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable serine/threonine-protein kinase 1 (EC 2.7.1.-).
 GN TC0575.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83560;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=MOPn / N199;
 RX MEDLINE=20150255; PubMed=10684935.
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MOPn and Chlamydia
 RL pneumoniae AR39.";
 CC Nucleic Acids Res. 28:1397-1406(2000).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE002326; AAF73573.1; -
 DR TIGR: TC0575; -
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 KW ATP-binding; Complete proteome.
 FT DOMAIN 4 296 PROTEIN KINASE.
 FT NP_BIND 10 18 ATP (BY SIMILARITY).
 FT BINDING 33 33 ATP (BY SIMILARITY).
 FT ACT_SITE 138 138 BY SIMILARITY.
 FT SEQUENCE 934 AA; 107498 MW; F08AEFD10A47E05C CRC64;
 SQ
 PKN1_CHIMU Length: 934 May 30, 2002 09:32 Type: P Check: 1219 ..

1 MORYELIRLI GRGGMGEVYL AHDKACGRV ALKTIREDLS DNPILRRKFL
 51 REAKIADILI HPGIVPYYSI CSDESVYYT MPEYEGFSIK HLLSWQKE
 101 ILSKELEERT SVKAFLPFD KICATVEYIH SKGLVLRDLK PDNILLGFG
 151 EYVYIDWQGA IFKAKELQO EKDEGRSSY GQKNICYSSM TVPKIVTTP
 201 DYMAPESLIG AEASEKTDYI ALGLILYQML TISFPYRRKK GRKLPEYDSI
 251 LSPIEAPYR EIPPSLSQIA MKAIANDPVQ RFSVQELRK ALQPHLGES
 301 EMTYDILST KDRKNWKYIE PILLSRFPV LASSPAQYN FMLSMEYNS
 351 SVRVECSYTR SSVQEGVGF FPPSKRDKG EFGYGLWF SSQNNELSVS
 401 LKNGIEIQG ESQGIIPQOS RFAISIEKSN NKITVEVDQI LEILHIDYLP
 451 SLGEREIIII QDLQISNIT ILSIGALRY SCLAYPADAL AEKLYDAQAR
 501 FYRKIRDSFP GRKESYEAOF RLAVTLITQI EEOGGLMQA LSTDLLHGS
 551 TGAPLEYLQK ALVYQNGSF VEETIRSLLA LKRYPOHPEI PRKDHLCFR
 601 LYDSLHKHRS EALVFMILLI WIAPEKIGLR EEPRTFLFH HRQOSTLFCR
 651 IDKTLQFKS SKMELFUSFW TGFTLFLPEL FQARADLQY QALIDIFYV
 701 CASGNKEVFS QFADLAIFV DEVVFPKSLH NORGEELYLF VQGLAALQNR
 751 EYRQAKFIS AVPPALQIYA LDLEPSLOAFI DEEVVFPSP LODIYNKNSA
 801 EDHKHLYVYM IQVSLMNQDL KQAYELLSKN PPODKGLIEY SEAFVLMGCV

851 LALGDRSAV KAHFSRCQF YGRSALICG IDDSDLDYLE GLVWMEKKT
 901 LFOSTFLIRC LHAPEKREY YROAYISMEN SFFG
 11AA_SEQUENCE 1.0
 ID PKN1.CHLPN STANDARD; PRT; 932 AA.
 AC 092986: 09K216;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable serine/threonine-protein kinase 1 (EC 2.7.1.-).
 GN CP00095 CP0679 from CPJ0095.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83538;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CM1029;
 RX MEDLINE=99206606; PubMed=10192388;
 RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickley E.K., Peterson J., Ullrich J., Berry K., Baas S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwyn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shirai T., Ishii K., Hattori M., Kubara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 from Japan and CML029 from USA";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE001595; AAD18248.1; ALT_INIT.
 CC EMBL: AE002226; AAF73695.1; ALT_INIT.
 CC EMBL: AP002345; BAA98305.1; ALT_INIT.
 CC TIGR: CP0679;
 CC InterPro: IPR000719; Euk_Pkinase.
 CC InterPro: IPR002290; Ser_thr_pkinase.
 CC InterPro: IPR001245; Tyr_Pkinase.
 CC Pfam: PF00069; Pkinase; 1.
 CC SMART: SM00220; S_TKc; 1.
 CC SMART: SM00219; TYRCK; 1.
 CC PROSITE: PS0107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
 CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 CC ATP-binding; Complete proteome.
 CC DOMAIN 4 291 PROTEIN_KINASE.
 CC NP_BIND 10 18 ATP (BY SIMILARITY).
 CC BINDING 33 33 ATP (BY SIMILARITY).
 CC ACT_SITE 138 138 BY SIMILARITY.

SQ SEQUENCE 932 AA; 107409 MW; 47AAAF87DF6CC5DF7 CRC64;
 PKN1.CHLPN Length: 932 May 30, 2002 09:32 Type: P Check: 8398
 1 MERYDIARII GKGGMGEVL AYDPCSRKV ALKKIREDLA ENPLKRRFL
 51 REARIADLI HPGVPPYTT YSEKDPYTT MPYIEGTYLK TLKSWQKE
 101 SLSELAERT SVGAFSLIFH KICCTIEYVH SRGILHRDLK PDNILLGLFS
 151 EAVILDMCAA VAGGEEDLL DLDVSEEVY SSRMTIPGRI VGPDPVAPE
 201 RLHGHPASKS TDIIYALGYV YQMLTISPPY RKKGKATVL DGGRISSPOE
 251 VAPREIPEP LSAVVMKMLA VDPQERYSSV TELKEDIESH LKSPKVTLT
 301 TALPPKSSS WKLNPEILLS KYFPMLEVP ASWSLAISSN IESFSEMRLE
 351 YTLKSKGLNE GRSILLPTE NALGDFPYOG YGFMLHKER TLSVSLVKN
 401 LEIQRCSQDL ESDKEFFLIA LEQHNHSLSL FVDGTTWLIH MNYLPSRSGR
 451 VAIIVDMED ILFDIGIFES SGLRVSCLA VPDAFLAEKL YDRALVYRR
 501 IAEFPGRKE GYARRRACI TVLEKASTDN NEOEPALAE EFSKLDGVA
 551 APLEYLGRAL VYORLOEYNE EIKSILLALK RYSPHEIFR LKHVVYRLH
 601 ESFPRDLRLA LVFMILVLEI AQQAITPGE EKLVLWLKDK SRATLPCLLD
 651 PTVELLSSK MELFLTSWGS FIPHLNSLFH RAMDOSDVRA LIELFYACD
 701 LHKQFLSSC IDIFKESLD QKATEIYEF SFEDIGALFL AIOISFNKED
 751 AEKIFVSDQ LSPILLYYIF DLFANRALLE SQGEAIFQAL DLIRSKVPE
 801 FYHYILRNHE IRAHLMCRNE KALSTIFENY TEKOLKDBQ ELVLYGCVL
 851 ALIIGAFAAK QHEDVCREDR IEPASILARN YNRLGLPKDA LSYOERELL
 901 RQKFLYFHL GNHDERDLQ TWYHILTEEF QL
 11AA_SEQUENCE 1.0
 ID PKN1.CHLPN STANDARD; PRT; 934 AA.
 AC 084303;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable serine/threonine-protein kinase 1 (EC 2.7.1.-).
 GN C1301.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D/OW-3/CX;
 RX MEDLINE=9900809; PubMed=9784136;
 RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.;
 RT "Genome sequence of an obligate intracellular pathogen of humans:
 RT Chlamydia trachomatis.";
 RL Science 282:754-759(1998).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR EMBL: AE001302; AAC67894.1; -.
 DR HSSP: Q06486; ICK1.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; Pkinase: 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP: 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM: 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST: 1.
 DR Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 KW ATP-binding; Complete proteome.
 FT DOMAIN 4 296 PROTEIN KINASE.
 FT NP_BIND 10 18 ATP (BY SIMILARITY).
 FT BINDING 33 33 ATP (BY SIMILARITY).
 FT ACT_SITE 138 138 BY SIMILARITY.
 SO SEQUENCE 934 AA; 107665 MW; A06EC6F4B025DADB CRC64;
 PKML_CHLTR Length: 934 May 30, 2002 09:32 Type: P Check: 4900 ..

1 MOREYLRLI GKGMGEVYL AHDKCSRRV ALKRIREDLS GMLLRQRL
 51 REAKIADLI HPGIVPYSI CSDGEAVYTT MPYEGFSIK SLKSVQKE
 101 VLSKLEKKT SVKSLPIPIF KICATVEYTH SKGVLRDLK PDNLLGLFG
 151 EVVIIDGAA IFKHARELKL EODDEAAVSF DERNICYSM TTPGKIYGP
 201 DYMAPESLG VEASEKTDIV ALGLILYQML TLAPFYRRKK GRKLSYRDV
 251 LPPLEMSYR EIPPSLSQIA MKAIAINPAD RPSISQELRQ ALQPYLOGDP
 301 EMTYKATLMA KEKSCWKYYD PILLSRYFPV LASSPAQYN FMLSVEISA
 351 STRVEYTVTK SAVHEGWLIL FLPSKEARG EFGCYGLMF SVQNHETLVS
 401 LKNGIEIQQK KSGEMISQOS RRAILIEKSD NRIAEVQA LFLIHIDYLP
 451 SLGNRLGVI QDLQMSNIA IESISGALRV SCLAVPAFL SEKLYDQAI
 501 FYRRIRDSFP GRKESYEAFQ RLGVTLLTQI EEGGDLTQA LSPFDYLGCG
 551 AGALEVEYIK ALVYQRNGSF VEIRICLLFA LKRYQHHEI PRLEDHLCFR
 601 LYDSLHKHRS EALVFMILLI WIAPEKISVR EEEFLRIITV HKQATLFCQ
 651 VDKAPLQFRS SKMELFLSFV TGFSLFLPEL FFRAGGLRDY QALADIVYA
 701 GVSNGREAFM QFSTALANVS DETTPESLIH NQKVALMFP VKGYEALRNK
 751 DYQAKKILLW KTFPTQLYA LDMFHIAFL DEETESFIDL LQALYDPASE
 801 EERDHILVYI IQTHLMNRDL ERAVYKLLNR FPLDEBLAEV SEAFILMGCV
 851 LALUGDRVYV KAHFSRCRYK YGKSALIGKC VDGDFIDYLD NLVMEKKMT
 901 LFQSYFLIRC LNESPRYER YRQAYLSMEN NFFD
 11AA_SEQUENCE 1.0 STANDARD; PRT; 459 AA.
 ID PUCC_RHOSH
 AC 002443;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein pucc.
 GN PUCC.
 OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Rhodobacter.
 OC NCBI_TaxID=1063;
 RN 11
 RP SEQUENCE FROM N.A. / 2.4.1 / NCIB 8253 / DSM 158;
 RC STRAIN-ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
 RX MEDLINE=93086425; PubMed=1453956;
 RA Gibson L.C., McGlynn P., Chaudhuri M., Hunter C.N.;

RT "A putative anaerobic coproporphyrinogen III oxidase in Rhodobacter
 RT sphaeroides. II. Analysis of a region of the genome encoding hemf and
 RT the puc operon";
 RL Mol. Microbiol. 6:3171-3186(1992).
 RN 121
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
 RX MEDLINE=20115911; PubMed=10648776;
 RA Choudhary M., Kaplan S.;
 RT "DNA sequence analysis of the photosynthesis region of Rhodobacter
 RT sphaeroides 2.4.1";
 RL Nucleic Acids Res. 28:862-867(2000).
 CC -1- FUNCTION: PUCC IS REQUIRED FOR HIGH-LEVEL TRANSCRIPTION OF THE
 CC PUC OPERON.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X68796; CA48701.1; -.
 DR EMBL: AF195122; AAF24245.1; -.
 DR PIR: S28025; S28025.
 KW Antenna complex; Transmembrane.
 FT TRANSMEM 31 51
 FT TRANSMEM 66 86 POTENTIAL.
 FT TRANSMEM 106 126 POTENTIAL.
 FT TRANSMEM 136 156 POTENTIAL.
 FT TRANSMEM 179 199 POTENTIAL.
 FT TRANSMEM 207 227 POTENTIAL.
 FT TRANSMEM 260 280 POTENTIAL.
 FT TRANSMEM 298 318 POTENTIAL.
 FT TRANSMEM 327 347 POTENTIAL.
 FT TRANSMEM 350 370 POTENTIAL.
 FT TRANSMEM 384 404 POTENTIAL.
 FT TRANSMEM 429 449 POTENTIAL.
 SO SEQUENCE 459 AA; 48973 MW; CFDD43326C9C07B1 CRC64;
 PUCC_RHOSH Length: 459 May 30, 2002 09:32 Type: P Check: 7732 ..

1 MSRIAEHLVR IGPREFPAD AASDQLPRK LLRLSLFOVA VGMAIVLVG
 51 TLNRMIYEL KVPASVYGM ISLPFLFAP RALIGFKSDT HVSALGKRV
 101 PMYIRGTAL WGGFALMPFA LYLGGQGYA EGQPFMLGVS SAALAFPLVG
 151 GGVHTIQTVG LALATDLAPR EDQPKVGLM YVLLISMIF ASIGFGLLD
 201 PYDAQLIKV ISGVAAVVF LNMIALMKME PRNRAFVYKP EKEPEGDHW
 251 REFSIRENAL HGLIVIGLGT LGFGADVIL EPGYGEVLSM TVAETTRLTA
 301 TPAGGGLVGF WLASWVLGRG FDLPRMAFLG AAAGLPGFPA IMGATEFTNV
 351 WFLIGTLIVV GFGGGLFSHG TLTATRLAPR KEQVGLAIGA WGAOVAATAAG
 401 VALAGAVLVR DILQMPDLS GYGPAPYVA VFALERGLFL LTMIVILPLL
 451 RSLAARRL
 11AA_SEQUENCE 1.0 STANDARD; PRT; 535 AA.
 ID TCE2_AVESSA
 AC P54411;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE T-complex protein 1, epsilon subunit (TCP-1-epsilon) (CCT-epsilon)
 DE (TCP-R36).
 OS Avena sativa (Oat).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Poace; Avena.
 OX NCBI_TaxID=4498;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. PEMI; TISSUE=Mesocotyl, and Coleoptile;
 RX MEDLINE=94085629; Pubmed=7903257;
 RA Elmann B., Krenz M., Mummert E., Schaefer E.;
 RT "Two Tcpi-1-related but highly divergent gene families exist in oat
 RT encoding proteins of assumed chaperone function.";
 RL FEBS Lett. 336:313-316(1993)
 CC -1- FUNCTION: MOLECULAR CHAPERONE; ASSIST THE FOLDING OF PROTEINS UPON
 CC ATP HYDROLYSIS. KNOWN TO PLAY A ROLE, IN VITRO, IN THE FOLDING OF
 CC ACTIN AND TUBULIN.
 CC -1- SUBUNIT: HETERO-OLIGOMERIC COMPLEX OF ABOUT 850 TO 900 KDA THAT
 CC FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE Tcpi-1 CHAPERONIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X75778; CAA53397.1; -
 DR HSSP: P48424; 1A6D.
 DR InterPro: IPR002423; TCPL_cpn60.
 DR InterPro: IPR002194; TCPL_1.
 DR Pfam: PF00118; cpn60_TCPL_1.
 DR PRINTS: PR00304; TCOMPLEXTCP.
 DR PROSITE: PS00750; TCPL_1; 1.
 DR PROSITE: PS00751; TCPL_2; 1.
 DR PROSITE: PS00995; TCPL_3; 1.
 DR Chaperone; ATP-binding, Multigene family.
 KM SEQUENCE 535 AA; 35034 MW; 02761226F5b17881 CRC64;
 SQ
 TCE2_AVESa Length: 535 May 30, 2002 09:32 Type: P Check: 4989 ..
 1 MALDFEDYWR PFILREGEK KSRLOGIDAQ KANIAACKSV ARLRLTSLG
 51 KMDKMLQSP DGDVITINDG ATELEMDVD NQIAKLMEV SRSQDYDIGD
 101 GTGVVVMAG SLLEQAELK ERGHIRIVA EGYEMASRIA VDHLESISTK
 151 YESATDIEP LVQTCMTTIS SKIVSRCKRA LAELAVAVL AVADLERKDV
 201 NDLIKVEKG VGKLEDETEL VOGITVYKDM SHQMPKRIE DAHIAITLCP
 251 FEPPKPKTKH KVDIDTVEKF QTLRQEQRY FDEWYORKCD VGATLVICQM
 301 GPDEANHL MLRELPAVRM VGVVLELELA IATGGRIVPR FQGLSTENKLG
 351 KAGLIVERKSF GTTKDRLMTI EKCAKSAVT IFINGKMM IETKRSIHD
 401 ALCAVRMLII NNSIYVGGG AEISCSIAVE AAADRHGVE QYAIRAFADA
 451 LDAIPLALAE NSGLPIIDTL TVVKSQHKKE NNSRCIGIDCN DVGTDNMKEQ
 501 NVEFTLIGKQ QOILLATQYV KMLIKIDVTI TPSEY
 11AA_SEQUENCE 1.0
 ID TRL2_HUMAN STANDARD; PRT; 1503 AA.
 AC 094759;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Long transient receptor potential channel 2 (LTPRC2) (Transient
 DE receptor potential channel 7) (TtrpC7).
 GN TRPC7 OR KNP3.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99026133; Pubmed=9806837;
 RA Nagamine K., Kudoh J., Minoshima S., Kawasaki K., Asakawa S., Ito F.,
 RA Shimizu N.;
 RT "Molecular cloning of a novel putative Ca²⁺ channel protein (TRPC7)
 RT highly expressed in brain.";
 RL Genomics 54:124-131(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20289799; Pubmed=10830953;
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
 RA Oniki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
 RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
 RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
 RA Schafke M., Schoen O., Desario A., Reichelt J., Kauer G., Blocker H.,
 RA Ramser J., Beck A., Klages S., Hennig S., Rietzelmann L., Dagand E.,
 RA Wehrmeyer S., Borzym K., Gardiner K., Nitzel D., Francis F.,
 RA Leherich H., Reinhardt R., Yaspo M.-L.;
 RT "The DNA sequence of human chromosome 21.";
 RL Nature 405:311-319(2000).
 CC -1- FUNCTION: MAY BE A CALCIUM CHANNEL.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE TRANSIENT RECEPTOR FAMILY. LTRPC
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB001535; BA34700.1; -
 DR EMBL: AF001754; BAA95563.1; -
 DR MIM: 603749; -
 DR InterPro: IPR002111; Cat_channel_TrpL.
 DR InterPro: IPR000636; Cation_chan_non_lig.
 DR InterPro: IPR000086; NUDIX_hydrolyase.
 DR InterPro: IPR002153; Trans_recep.
 DR Pfam: PF00520; Ion_trans. 1.
 DR Pfam: PF00293; NUDIX. 1.
 DR PRINTS: PR01097; TRANSRECEPTR.
 KM Ionic channel, Transmembrane, Ion transport, Calcium channel.
 FT DOMAIN 1 752
 FT TRANSMEM 753 773
 FT DOMAIN 774 795
 FT TRANSMEM 796 816
 FT DOMAIN 817 820
 FT TRANSMEM 821 841
 FT DOMAIN 842 896
 FT TRANSMEM 897 917
 FT DOMAIN 918 936
 FT TRANSMEM 937 957
 FT DOMAIN 958 1025
 FT TRANSMEM 1026 1046
 FT DOMAIN 1047 1503
 FT TRANSMEM 1047 1189
 FT CONFLICT 1189 1189
 SQ SEQUENCE 1503 AA; 171225 MW; AD329AE79F1A71B5 CRC64;
 TRL2_HUMAN Length: 1503 May 30, 2002 09:32 Type: P Check: 5055 ..

1 MEPSALRRAG SEQEGEGEGL PRRVTDLGMV SNLRSSNSL FKSRLQCPF
 51 GNNDKQESLS SWIPENIKKK ECYVEVSSK LSDACKVVCQ CGYHEQHLE
 101 EATPRHTQGS TQMDPKKHVQ EMPDAGFQI VFTGISQVKK KYAVVSDDP
 151 SSVIYHMTQ HWGLDVPNLL ISVTGAKNF NMRPLKLSIF RRLGVKAQT
 201 TGMNITGGS HTGVMKQVGE AVADFSLSSS YKEGELITIG VATGTVHRR
 251 EGLHPGSGF PAEYIIDEQ OGNTCLDGN HSHFLYVDG THGYGYEIP
 301 LRTLEKEIS EQTKERGVA IKIPVCYVL EGGPGTLMTI DNATNGTSPC
 351 VVVEGSGRVA DVIAOVANLP VSDITSLIQ OKLSVFEQEM FETTESRIV
 401 EMTKKIQDIY RRRLTVFR EGGDGGQDAD VALIDALLKA SRSDHGHGE
 451 NMDHQLKLA VAMNRVDIARS EIFMDQMOK PSDLHPMTA ALISNKBEFV
 501 KLFENGVOQL KEFYVMDTL YLYENLDPSG LFHSKLQKVL VEDDERACA
 551 PAARLQMHV VAQVLRLLG DFTQPLTPRP RHNDRLRLLL PVRYVKNVQ
 601 GVSLSILYKR SSGHVPFTMD PIRDLTIMAI VQNRRELQGI IMAOSODCIA
 651 AALACSKILK ELSKEEDTD SSEMLALAB EYEHRAIGVF TECYRKDEER
 701 AOKLITRSE AMGRTTCLQ ALEAKMKFV SHGIGQAFIT KVMGQLSVD
 751 NGLMRVTLGM IAPPLLTGL ISFERKRLQD VGTAPAARARA FETPAVVVFH
 801 LNTLSYFAFL CLFAYVLMVD FQVPSPMCEC AIYLMFLSYL CEEKRLQLEYD
 851 PDEGIMAKKA ALYFSDPMNK LDVGALLFV AGLTRILIPA TLTYGRVILS
 901 LDFILFCLRL MHITTSKTL GPKIIVKRM MKDVEFFELF LAVWVSGFV
 951 AKQAILIHNE RAVDWLFRGA VYHSYLTIFG QIPGYIDGVN FNPBSPNG
 1001 TDPYKPKRPE SDATQORPAF PEMLYVLLIC LYLETNILL LNLIAFNV
 1051 TFOOVQEHNT QIMKFORHDL IEEYHGRPA PPPFILLSHL QLTFRVVLK
 1101 TPAKRHKQLK NKLKNEEA LLSMEIYLKE NYLQNRQFQ KQREOKIED
 1151 ISNVDAVMD LLDLDPLKRS GSMEOQLASL EEOVAQTARA LHMIVTRLRA
 1201 SGFSSEADVP TLASQKAAEE PAEPGGRKK TEEPDSYHV NARHLLYPNC
 1251 PVRPEPVNE KVPWETEFLL YDPPFYTAR KDAAMDPNG DTLPELSTIQ
 1301 YNVVDGLDR RSFHGPTYQ AGPLPMPGR TGLRGSGLS CFGNHLTP
 1351 MYTRMRNED GAICRSIKK MLEVLYVLP LSEHWALPGG SREPGMLPR
 1401 KIKRILROEH WPSFENLKC GMEVYGYMD DPRNTDNAMI ETVAVSVHFQ
 1451 DQNDVELNRL NSNLHACDSG ASIRMOYVDR RIIPLANKIT LLOKAAEFEG
 1501 AHY

11AA_SEQUENCE 1.0
 ID VES_RHPV1 STANDARD; PRT; 157 AA.
 AC P24834;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Probable E5 protein.
 GN E5.
 OS Rhesus papillomavirus type 1 (Rbpy 1).
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 CC Papillomavirus.

OX NCBI_TaxID=10570;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91135018; PubMed=1847267;
 RA Ostrow R.S., Ladresh K.V., Faras A.J.;
 RT "Characterization of the complete Rbpy 1 genomic sequence and an
 RT integration locus from a metastatic tumor";
 RL Virology 181:424-429(1991).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M60184; AAA79315.1; ALT_SEQ.
 DR EMBL; M60184; AAA79316.1; ALT_SEQ.
 DR PIR; F38503; W5MLR1.
 KW Early protein.
 SQ SEQUENCE 157 AA: 17398 MM; AC7AA67158844686 CRC64;

VES_RHPV1 Length: 157 May 30, 2002 09:32 Type: P Check: 8743 ..

1 MVMCIGTQMS HKPVHTLNS IQVLCRANCC CYACKRPPEC CFMLCFCFCF
 51 CLALCFVHLL SRCEVCPVPC LSVAAVAVVL GVHSEPVCSF WSVVLEFPNP
 101 VAFDTPACQ CGLQNDVNT AHRHVIISYF AIVAVNIYFV LALLVGAAFK
 151 ATSRART

11AA_SEQUENCE 1.0
 ID VP41_BPAPS STANDARD; PRT; 460 AA.
 AC Q9T1Q7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative protein p41.
 GN 41.
 OS Bacteriophage APSE-1.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae.
 OX NCBI_TaxID=106199;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99420383; PubMed=10489345;
 RA van der Wilk F., Dullemans A.M., Verbeek M., van den Heuvel J.F.J.M.;
 RT "Isolation and characterization of APSE-1, a bacteriophage infecting
 RT the secondary endosymbiont of acyrthosiphon pisum.";
 RL Virology 262:104-113(1999).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF157835; AAF03984.1; -.
 DR KW Hypothetical protein.
 SQ SEQUENCE 460 AA: 51920 MM; 9572842E1FFDE31 CRC64;

VP41_BPAPS Length: 460 May 30, 2002 09:32 Type: P Check: 4265 ..

1 MKTFPRPYQ HLIINHLDI KRSNVWAGM MKTAATLTA LBNLYLSGSE
 51 TKPLVLAPL RVAQSTWPE ALKWSHLRNI EVQPIIGSAK ARIALAKNTH
 101 ASVFYVANDN LVMVLVDLGD TMEFGITIND ESTRLAKSRL RKGGKRIAL
 151 AKIAHKSVRH WNLVTGTPSP NGIMDIMGQA WFDQGERLIG RTYNAPTFRW

201 FKRLQPLGCO WSRFPLGFA HLQPLAUSD VTLSDAADW FDIDEPIHNV
251 INVELPAKAR ANYHAMEKEI FLELGESAIE ALNAAATIK ILQIASGAIY
301 SDOARNTEI HDKIQALES IYNSGCTPV LVAYHWKHD LERLKAPKPK
351 KNLDPANRTL TDMNNGRIPL LEAHPASGCH GLNIQDGGNI LVFPHWMDL
401 EGYOIIERI GPTROAAGH NRPEIHHIV AKDTLDEVVM ERNRSKRAIQ
451 DLLEAMKRR
11AA_SEQUENCE 1.0
ID Y306_MYCPN STANDARD; PRT; 395 AA.
AC P75343;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG306 homolog (A05_orf395).
GN MPN435 OR MP406.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID-2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE-97105885; Pubmed-8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
Hermann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
DR EMBL: AE000039; AAB96054.1;
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 15 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 86 106 POTENTIAL.
FT TRANSMEM 131 151 POTENTIAL.
FT TRANSMEM 175 195 POTENTIAL.
FT TRANSMEM 254 274 POTENTIAL.
FT TRANSMEM 298 318 POTENTIAL.
FT TRANSMEM 348 368 POTENTIAL.
SQ SEQUENCE 395 AA; 45633 MW; 9E9C1ABD25B89B16 CRC64;
Y306_MYCPN Length: 395 May 30, 2002 09:32 Type: P Check: 3483 ..
1 MSPPKSASF TRKNIAFSF FIAFLVVSV LVTFELDIR TGDVKTINT
51 INRTNPMIL LIVLIVVTL AMNIIIMMV ARRCFHAPW WENVLACVY
101 QFOIYVPLS LGDPPRLYV FIKKMKKOT AVLLYSTGA FYNLAQAALT
151 WPSFEVLSON VALLQNHG FVSYWSFAG MIEDVVAAIL FPIAASKRM
201 HVLISGVNQ FRKMKRPL TREQIYQRFI DKAEPNLYG LEIKRIGLTI
251 FKLILANLIA VVGYSVFAY FAIYKKNAT NNVIDQYSTA DIFNTINIAI
301 TASNFIPLPS GEGATQFVMT SFLNAFKSAV GIESOVKGV FLWRFLSVYI
351 PALFSLCLFI GMVYGVVIE KHPKRVLPV NLINHHWNN KKLHN

11AA_SEQUENCE 1.0
ID Y448_MYCPN STANDARD; PRT; 518 AA.
AC P75066;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN048 (D05_orf518).
GN MPN048 OR MP106.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID-2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE-97105885; Pubmed-8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
Hermann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SIMILARITY: BELONGS TO THE MG032 / MG096 / MG288 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
DR EMBL: AE000013; AAB95754.1;
DR InterPro: IPR004306; DUF237.
DR InterPro: IPR004319; DUF240.
DR Pfam: PF03072; DUF237; 1.
DR Pfam: PF03086; DUF240; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 518 AA; 58873 MW; A0D753FC7DC22EC CRC64;
Y448_MYCPN Length: 518 May 30, 2002 09:32 Type: P Check: 4533 ..
1 MNEIQTIVNE ANNIPVEAP KVPKPEKLF GFVNGGFPP LNPKGINTND
51 NVASLEFQYS LKQASLKDFD ILLEKKNIV LEHKRVYNA LOHFPEITYV
101 GTGEINILQF ALQASTNFS SLEELQASFS KTGDNILTAQL FMKPYTKLV
151 SGENDLTHIA QTAIGSLFD SRVDSASII NSEATLKTAE ATPTTQVLP
201 FKAEREKALA IKKAEERIK KELEOKRRQ EELSKOORDK EALQKSLNMF
251 QEFISYWTGQ GHDVKOEQF IQALEAFST NMNVEFNILL AGRSRAIQTY
301 YKQSKAQSQ NAKIAGEKG IQPKSGPL DGIFMSPLR GNLTGAHND
351 LKLKVEVN TOGKDAQND KKASINQAK QNNEPFRQV PWFSEFEVL
401 KYEGSYGLPV GARFLNFGS LGIPDMKGE MSYFVLIDGK TPQWIDKRD
451 YPSGLPEFEK NOLKTPPHK EHVHVENKQF MEKIKSNLH NIELATGARK
501 PPAVDLASYL HYLILNHK
11AA_SEQUENCE 1.0
ID YC03_KLEPN STANDARD; PRT; 504 AA.
AC Q48449;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 55.8 kDa protein in CPS region (ORF3).
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;


```

OC Klebsiella.
OX NCBI_TaxID=573;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHEDID;
RX MEDLINE=95204345; PubMed=7896702;
RA Atakawa Y., MachimotoYankun R., Nagatsuka T., Ito H., Kato N.,
RA Ohta M.;
RT "Genomic organization of the Klebsiella pneumoniae cps region
RT responsible for serotype K2 capsular polysaccharide synthesis in the
RT virulent strain Chedd."
RT J. Bacteriol. 177:1786-1796(1995).
RL -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D21242; BAA04774.1; -
KM Hypothetical protein
SQ SEQUENCE 504 AA; 55782 MW; AD887595CFDFPAB CRC64;

YC03_KLEPN Length: 504 May 30, 2002 09:32 Type: P Check: 2748 ..

1 MPRALSTIVE PPSFLRYDSG AENSCTMIK IARIAVTLGL LSLGQAVV
51 AGLVNDNDL RNDLAWLSDR GYIHLSLSTW PLSEIEISRA LKAKKSYSS
101 EGVVLARINQ RLALKADFR VTGYSTDPK GPPOGFCOTQ PADNSLGLAF
151 NNSGEMDVH LOGNVEGGER ISNGSRFNAN GAYGAVKFWN QWLSFGQVPO
201 WNGPGYIGSL INGDMRPMT GFLMORACDA APETWMLRW GPWOYQISAS
251 QMNOYTAAPH AKIIGRTFF TPFOSLLELGA SRIMQMGEG RPOFSFQMD
301 GFTGHDTGT DNEPQNLAG FDFKFLKLEPT LGMPVSFYQ MWGEDSGYL
351 PSANMPLGCI EGHGHWKDA VMWYEADHT RTNMSRNYS YTHIHKDGY
401 YQGYPLGDA MGDGQLFAG KVELITENNO RMSTRLAYAK VNPKDOSINK
451 AFPHSDTLKG VOLGWSGDYV QSVRLNTSLM YTNANNSDSD DVGASAGIEI
501 PPSL

11AA_SEQUENCE 1.0
ID YC34_MYCPN STANDARD; PRT; 417 AA.
AC P75537;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN234 (G07_orf417).
GN MPN234 OR MP597.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfeich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SIMILARITY: BELONGS TO THE MG032 / MG096 / MG288 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000058; AAB96245.1; -
DR InterPro: IPR004306; DUF237.
DR InterPro: IPR004319; DUF240.
DR Pfam: PF03072; DUF237; 1.
DR Pfam: PF03086; DUF240; 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 417 AA; 47650 MW; 8880670173D369C9 CRC64;

YC34_MYCPN Length: 417 May 30, 2002 09:32 Type: P Check: 7095 ..

1 MGFHAYYQNN DTKANDFAL NISTINPATL QELANSFDLQ GSDLTAGLFF
51 KYSVNLITSG TNDLTIART ALGENIIOKQ VSLTQSIIRK RLEBAKTQYK
101 QDIAPFAKE RQALAQHUK EIEEAKQRAE QLEKQGEAE KRQEEYKNV
151 AETQGFNDLS TSAQKREYV LKQKQVTKK VELIQALKSS FFRNQRFTN
201 FLIAGFRTAI DWYNOEKNN TTAKNNAFCK NGIOFPVAGF QGIYMSQWLR
251 DELSGKTDIK LNKLSLSYON ENKNSIMWN KQKIEIKQV KPNYSFEIN
301 LKTYGSYVNS LMYLGAIGA GIPTSWSGTM DMKEPIVDGL DSGIVTKQDY
351 PGSKFEETED KLMFTLVKQ QIKVKEGCFM NLKQGSIDN LDRTGTTRK
401 PYVDLASYLH FVILTAK

11AA_SEQUENCE 1.0
ID YCF2_MESVI STANDARD; PRT; 890 AA.
AC Q9MUP8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 103.9 kDa protein ycf2 (RF2).
GN YCF2.
OS Mesostigma viride.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;
OC Mesostigmatales; Mesostigmataceae; Mesostigma.
OX NCBI_TaxID=41882;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIES-296;
RX MEDLINE=20150907; PubMed=10688199;
RA Lemieux C., Oltis C., Turmel M.;
RT "Ancestral chloroplast genome in Mesostigma viride reveals an early
RT branch of green plant evolution."
RL Nature 403:649-652(2000).
CC -1- SIMILARITY: BELONGS TO THE YCF2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF166114; AAF43852.1; -
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003960; AAA_sub.
DR InterPro: IPR003959; AAA_subfam.
DR Pfam: PF00004; AAA_1.
DR SMART: SM00382; AAA_1.
DR PROSITE: PS00674; AAA_1.

```

KW Chloroplast: Hypothetical protein.
SQ SEQUENCE 890 AA; 103935 MW; 43CAEE991AF2C4B CRC64;

YCF2_MESV1 Length: 890 May 30, 2002 09:32 Type: P Check: 5630 ..

```

1 MOQHMSLSN MIKEVEVEF HVPALLIPTR YREIYREIDR APRLLQSFY
51 TIFKVNIT EKYVTIDVY YSSVEDPLGL OSPHSRMAQ NDEOROKA
101 RIISQVYRK NSFIRITSNK PFRVLITSLF VLFYFQSGV FQSPYFESQ
151 INGLFKKMN NINQLEINST NTCIIKITDL LTHQSKELL GNSLPINSL
201 RAKHKSISD IKTFMPEK WYLPGLGL ENSEQITSYR NPDICULTN
251 SSNIMKIKN RTLLDPKIDK YPLVLKKN ILFFHKLPIQ FFLPYIVLR
301 IMLAPFVLM WSYQNSEEK ENIKNNLKN HDIEISTIQ FVAKATFERD
351 IGMESLQOE LATVAFLLKQ KNSNSYPMG YLFGAPPGTG KITMAKAMST
401 EAETPYLYVE GSGFRCREG VANARVDLJF KQIQNISPCI LYIDEIDSLA
451 ERREANKOL EQLKITGDSI EGSNINIDQK PSDTVLMQFL IYMDGYKKRN
501 DLIIGATNR IETLDAIMR PGRFPRQIVF SPFFFEERKD ILRFILRNK
551 ALVDDTKTM MAERSTGLNG CDLRLLADNI LLSALESRN OOKTIPVINE
601 DFEDRALERV SRIRHISNV ELAFGYDFY RTAVHEAGIA LIHTLLPECR
651 PYSVAKLEPK PLNDRYLEIE RENLKVPSDD IISTNNIDYF VOKIVGLLAG
701 RAASESLDFE YPGQISTYLN KTYDPNIOGA YNIAHIVEF GLLDSTVTVI
751 HYNSENNEN IKDPEYTKIN DLIQNKVTLK TNRELKRSQ ASILDFNEF
801 WYEDDYWQF DFIKKOYIS NISSRLDME IVSLIHLFQ YTFDFLKSNE
851 QLDHLASLV LNKSTISQKE IHLVNSYGI KIPTKTKAM

11AA_SEQUENCE 1.0
ID YCLF_BACSU STANDARD; PRT; 492 AA.
AC P94408;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical transporter yclF.
GN YCLF.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97124189; PubMed=8969502;
RA Yamane K., Kumano M., Kurita K.;
RT "The 25 degrees-36 degrees region of the Bacillus subtilis chromosome:
RT determination of the sequence of a 146 kb segment and identification
RT of 113 genes."
RL Microbiology 142:3047-3056(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: BELONGS TO THE PFR2 FAMILY OF TRANSPORTERS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: D50453; BAA09000.1; ..
DR EMBL: Z99106; CAB12175.1; ..
DR Subtilist; BG12027; YCLF.
DR InterPro: IPR000109; PTR2.
DR Pfam: PF00854; PTR2; 1.
DR PROSITE: PS01022; PTR2_1; 1.
DR PROSITE: PS01023; PTR2_2; 1.
KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT TRANSMEM 67 87 POTENTIAL.
FT TRANSMEM 88 108 POTENTIAL.
FT TRANSMEM 110 130 POTENTIAL.
FT TRANSMEM 160 180 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 232 252 POTENTIAL.
FT TRANSMEM 255 275 POTENTIAL.
FT TRANSMEM 294 314 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 367 387 POTENTIAL.
FT TRANSMEM 392 412 POTENTIAL.
FT TRANSMEM 434 454 POTENTIAL.
FT TRANSMEM 464 484 POTENTIAL.
SQ SEQUENCE 492 AA; 53275 MW; 5042A6590D30E03 CRC64;

```

YCLF_BACSU Length: 492 May 30, 2002 09:32 Type: P Check: 9486 ..

```

1 MASIDNESII KSVPOKGFEG HPRGLTLEF TEMWERSYI GMRAILLYLL
51 YTFVNGGLG FDKGTAVAIM SIYGLVYMS TIIGWLADR VEGTANTVEY
101 GGIFIMNGHI ALAYPSSIA FYISWVLIIV GTGLKRNVS SVYGDLYTNE
151 DPRDSEFSI FTMGINLGL LAPLIVTGL QKYNHILFG MAAVGMLGL
201 IVEPLTRKKN LGIAGSNVPN PLSKSAIGT GIGVIVAIA VIISVOTGV
251 TIRFIDLVS IIGILIPYI FIIMTSKKA DKTKSLAA YPLFIGAAM
301 FMAIOEGAT ILAVYADERI RUSLGFELQ SSWFOSLNP FYVIEAPFA
351 LWMKLGKRO PSTPVKFSIG IILAGLSPII MVFPAMQKE ALVSPMLVL
401 SPFLVVLGEL CLSPVGLSVT TKLAPAFSA QTSMMWLTN AAOALNAQV
451 AGLEFKIPEP MYFGTIGLIS IVLGGILLLL SPVAKRAMG VL

11AA_SEQUENCE 1.0
ID YF11_MYCPN STANDARD; PRT; 260 AA.
AC P75275;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN511 (F04_orf260V).
GN MPN511 OR MP331.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscates;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfeich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RT Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SIMILARITY: BELONGS TO THE MG032 / MG096 / MG268 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AE000031: AAB95979.1: -

DR InterPro: IPR0004306: DUF237.

DR Pfam: PF03072: DUF237.1.

KM Hypothetical protein; Complete proteome.

SQ SEQUENCE 260 AA; 30422 MW; 85B7FD1D955FB6E CRC64;

YF11_MYCPN Length: 260 May 30, 2002 09:32 Type: P Check: 3241 ..

1 MQAFKEYWQK QKKDVTDKQ LLEALKLSFA KEQNKTFAPL IKNFQDCISN

51 YYPNDQEDQS EAAKTAFGTQ GIAFPQSGLK GIFMSEWLK OLGEKAKINL

101 DIKSLKVTDS KISPTIKMK DGIKRNQDK PYNFRFEIDI EYQGNKYLW

151 LEAIIAKFSG IPGEMKGLN LKFIYDGLS WEIVOKPDYP GSLPFDQK

201 QQLFKLHW EKITVQEPF MELIKSNLH NLELTESTK PPVVDLASYL

251 HYOLKLINQ

11AA_SEQUENCE 1.0 STANDARD: PRT; 1649 AA.

ID YG4_SCHPO

AC 060179;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical 186.4 kDa protein C33E.04c in chromosome II.

GN SPCC23E6.04c.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomyces.

OX NCBI_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RA Wood V., Rajandream M.A., Barrell B.G., Pohl T.;

RU Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE BAP28 FAMILY.

CC -1- SIMILARITY: CONTAINS 1 HEAT REPEAT.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AL023287: CA18872.1: -

DR InterPro: IPR000357: HEAT_repeat.

DR PROSITE: PS50077: HEAT_REPEAT: 1.

KM Hypothetical protein.

FT REPEAT 1608 1646 HEAT.

SQ SEQUENCE 1649 AA; 186409 MW; 6BAFDD373125216D9 CRC64;

YG4_SCHPO Length: 1649 May 30, 2002 09:32 Type: P Check: 5030 ..

1 MASSLOKOLK NIOSNNVLKI NKIRAPSL YDPKVAAMD LEEIYVTVAS

51 GFHELAVHEP RLTYFEKTL GEGSVQVDVY LNRTENEKI DLCEVOILRL

101 LAPFTEENA LKYLEMLIR FSIHEVSD E FILSLPHD HPFARILGC

151 SKPSRPLF LENAIPKPV LSRADIVHAL SRDKEFFAM AQFYONTAES

201 HNMTPELARF WAGTMEVLV AMHSNEDN VLDRFELRV SYAVSYSSI

251 DFOIAGFMLL SSIASLPLS PSIIPLVSA ITDRLSFDNI KPALICVGH

301 LQFCSSPEFD HEQLEKLESF GASSLIEIS QHRUDEEPV SYWVSLIKSR

351 KQKDKRLIS LDQTSISQIR VTHEQAKFL SVTPVNDQK ALQSVRRILD

401 SVIQPERKEG KLDNLINTLQ DKKSSTFSK KDEVLLAKI SEIDSQTSFE

451 QCLAYADSA DLDSSVEISL LSKFGDKIPF LLECIANGSE RIILSLIEL

501 KRTIEKNQV DYQIILPVYL YSLQSKDTEV RSRALNLIIT FLELRNEMLE

551 FSIITGMDN DNKNLRWLP VETKYICSDL LIDRSSEIGL DGTLYFSYIP

601 ERLFEKRPK NASKEIAVTS FLSSHAQSK ISNVVLLIE ILTFVHGKVE

651 DAKQIILPR LEQLEEFNSE KFKTVAKREV EALVCFNHT SFTSLSTLS

701 SNIVLSQAI CRIVEIOSHL KDPQRLFPK AVISODEOPH YYVDVLSIK

751 IPDVFKKLI GSVRLVKEKN PAIAKKRRID SHIPGDVQR LTRLLELLET

801 KNAASYPKLA SPLFEVLNSV IALKEDIVSS NVLLQLLGL LYEMIGASPI

851 TELSPSIRD TLVGCIRSTN NPQIQNKALL LVSALANAP EAVLHGVPPI

901 FTFMGSTVLS RDAFISIHVI EQTVKTIVISA LIRLGKEDS SLVSCFVNA

951 FPHIQHRL RLTVLVLOTI GSNRFLSVYL IQFAEKMLA KSTNVVAIHD

1001 PCLTVQSPS VADRIGSINQ CSRFCLKSLE EDSNDSNCK AVSLIKIDEL

1051 PMVDVLATIG SLRVKVLLEI SLVSKAKNFA FDLAKIMENS VDSEVEIQAG

1101 LEFSIKLITL LSQSSNME LGHVVALRS VYHLHPNEL CTYVGLKLLHD

1151 ERALLRRKAL STVOQVOQG SKVSALTALI PVTYINISY SDEETOLAM

1201 DCLAVMARF SASPELFISP IEVVSQPYGL KNSARDVQVS AIVCITVLTN

1251 TLAARILPYL ADIVNSLSI LDDARKDPEG DLEELACFSM MIDPFKVLPE

1301 FSSSIVEPTI KCAIASDRAE EMDAIGELF ETIANFIPPR LMKISIFAM

1351 PECARLAGSTA ALRLLELIEL ALONSSRSAL GTVYSIEFK FLDSFDSHRS

1401 LLFARDVQNV ETQAVNFKL FVAKLSDTTF RLFLHLHSM ALDEDYERDP

1451 SGIVSRQTFE YNFTLTPLDT LKSIYNYIA YVLDPTIEL SSKDTNSEVR

1501 HLVNSLSVA FENDTEEFWM VPARFGKISP VLIEQIYAP LDDKVLIVKA

1551 IVELASVASS SDNFRSMNTQ LLQYLSSNI NARLAIQIQ TQYGRIGEN

1601 WISTLPQSV FIAELMEDDD DOVETATIEL VRIIDRLGE NESLDYILT

11AA_SEQUENCE 1.0 STANDARD: PRT; 518 AA.

ID YL92_YEAST

AC P18634;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Hypothetical 59.8 kDa protein in SSRI-Arp10 intergenic region.

GN YLR392C OR I8084.13.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C / AB972;

RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D., Johnston L., Langston Y., Latreille P., Le T.,

RA Maridis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
RA Rifkin L., Riles L., Raich A., Trevasakis E., Vignati D.,
RA Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.,
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-210 FROM N.A.
RX MEDLINE=90277691; PubMed=2141026;
RA Ackerman S.H., Tzagoloff A.,
RT "Atp 10, a yeast nuclear gene required for the assembly of the
RT mitochondrial F1-F0 complex.";
RL J. Biol. Chem. 265:9952-9959(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U19729; AAB82352.1; -;
DR EMBL: J05463; AAB05631.1; -;
DR PIR: C36445; C36445.
DR SGD: S0004384; YLR392C.
KW Hypothetical protein.
SQ SEQUENCE 518 AA; 59758 MW; 65399EA2FFC2C23 CRC64;
..
Y192_YEAST Length: 518 May 30, 2002 09:32 Type: P Check: 3922 ..
1 MAPKISISLN PRNGEFYSS NDQMSGIVSL QLTAKLSIRK ISYLKGFSE
51 TLKRIQDEYM FOONGMMMPG QDNKSFHLM KFEORVPPD NVNNAIDGSS
101 KPRVKRPGSY NYSFOQDFKP RKPECLKNHT AKTYAEVTRS NARLPPFENS
151 HMOEFNKIDN LDLYEYFSGK VIYWOVOLE LGRSSWFKP FHKLIREIET
201 FEFIPPEKDL IIEPDEDNE ELNAPSNSR GNSWVTNNEF FNSNLKVPSS
251 KDYKVVNGVG YIKSDRFSQ ANSILIEGND IRSRPVSSVT STROSTRLVN
301 GMKVFSPSTY MGLPDGESHM RLEVRSDLK QIYRKDLFR SGSONDKVY
351 VVMGNTIASL SKMQLPPLK QLNLETTY LSOSIANGNY SSKLLEIDL
401 NQKSNKPL DLNREIREND GSMFECELRK KDHPILKLV FNEEDYRHG
451 NRLYSEKTCY IKRTESLQL IEMGIGIRK QSEVNIIDPVQ IFCQVREHVE
501 AEALPRVVP PYTEKAS
11AA_SEQUENCE 1.0
ID YXIK_BACSU STANDARD; PRT; 153 AA.
AC P42302;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yxik.
GN YXIK OR S58B.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=97124196; PubMed=8969509;
RA Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA Miwa Y., Fujita Y.;
RT "Sequencing of a 65 kb region of the Bacillus subtilis genome
RT containing the ltc and cel loci, and creation of a 177 kb contig
RT covering the gnt-sacxy region.";
RL Microbiology 142:3113-3123(1996).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D83026; BA11690.1; -;
DR EMBL: Z99124; CAB15950.1; -;
DR Sublist: BG1140; YXIK.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 153 AA; 17649 MW; CEC1799C0BF2A4DC CRC64;
..
YXIK_BACSU Length: 153 May 30, 2002 09:32 Type: P Check: 7683 ..
1 MEITSISIG NUDMILKPD QYVMACELED AESPYRRWAG LAYDRIMIOV
51 ITTGSFIEDL SEYFEGHAYK VTKLAKREPH FQSILOEADR DIADFLFLA
101 SINDVEFLIT DPQDRKSYFS ESKLOCLNDS GERITWEYD AVDIYIMIGE
151 SYK
11AA_SEQUENCE 1.0
ID Q9X4B6 PRELIMINARY; PRT; 504 AA.
AC Q9X4B6;
DT 01-NOV-1999 (Trembl,rel. 12, Created)
DT 01-NOV-1999 (Trembl,rel. 12, Last sequence update)
DT 01-DEC-2001 (Trembl,rel. 19, Last annotation update)
DE K30 CAPSULE BIOSYNTHESIS CLUSTER, PARTIAL SEQUENCE.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E69;
RX MEDLINE=99217011; PubMed=10200954;
RA Drummelsmith J., Whitfield C.;
RT "Gene products required for surface expression of the capsular form of
RT the group 1 k antigen in Escherichia coli (O9a:K30).";
RL Mol. Microbiol. 31:1321-1332(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=E69;
RA Drummelsmith J., Whitfield C.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF104912; AAD21561.1; -;
SQ SEQUENCE 504 AA; 55715 MW; 7A1E0A299A1DB841 CRC64;
..
Q9X4B6 Length: 504 May 30, 2002 09:32 Type: P Check: 2369 ..
1 MPRALSTLVE PSFSLRCDGSG AENGSCYMIK IARIAYVLGL LSLGQAAYA
51 AGLVNDNDL RNDLAWLSDR GVIHLSLSTW PLSGEETARA LKRAKPSYSS
101 EOYVLARINQ RLSALKADER VTGYSTDPQ GTPQGFQOTQ PADNSGLAF
151 NNSEEMDVH LGCNVNGGER ISNGSRFMAN GAYCAVAFWN QMLSFQOYD
201 WMGPGYEGSL IRGDANRPMT GFLMORAEQA APETMWLRWV GPWOYOISAS
251 QMNOYTAVPH AKIIGRTEF TPQSLLEGA SRIMQMGEG RPQSFSEMD
301 GFTCHDWTG DNEPQNOLAG EDFKFKLEPT LGWVSYEQQ MWGEDSGYL
351 PSANMFLGV EGHGNGKDT VMVYEADT RTNNSRTNYS YTHHIYKDG
401 YQOGYPLGDA MGGDQQLING KVELITEDNQ RWSRTRYAK VNPENQSIK
451 AFPHADTLKG LQUGWSGDYV OSVRLNTSLW YTNANNSDSD DVGASAGIEI

501 PFSL

11AA_SEQUENCE 1.0
 ID 091E1 PRELIMINARY; PRT; 70 AA.
 AC 091E1;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE SC3D11.21 PROTEIN.
 GN SC3D11.21
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Brown S.P., Harris D.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Bentley S.D., Parkhill J., Barrett B.G., Rajandream M.A.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denaplatte D., Eichner A., Cullum J.,
 RA Khashash H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL157916; CAB76018.1; -.
 DR HSSP; P02593; ICDM.
 DR InterPro; IPR002048; EF-hand.
 DR PROSITE; PS00018; EF_HAND.
 SO SEQUENCE 70 AA; 7742 MW; 82A559BF116480 CRC64;

091E1 Length: 70 May 30, 2002 09:32 Type: P Check: 7894 ..

1 MADIEARKQ FERIDTDGGS LITPAEFKTA LAQGGDMWNT ESVAEATIG

51 RDLDDKQLS FDEFWAHLNK

11AA_SEQUENCE 1.0
 ID P9529 PRELIMINARY; PRT; 492 AA.
 AC P9529;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE TDNQ PROTEIN (EC 6.3.1.2).
 GN TDNQ.
 OS Pseudomonas putida.
 OC Plasmid pTDN1.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97144524; PubMed=8990291;
 RA Fukumori F., Saint C.P.;
 RT "Nucleotide sequences and regulatory analysis of genes involved in
 RT conversion of aniline to catechol in Pseudomonas putida UCC22
 RT (pTDN1).";
 RL J. Bacteriol. 179:399-408(1997).
 CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) - ADP + GLUTAMINE +
 CC ORTHOPHOSPHATE.
 CC SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
 DR EMBL; D85415; BAAL2805.1; -.
 DR HSSP; P06201; ILGR.
 DR InterPro; IPR001637; GlnA_adenyltn.

DR InterPro; IPR001691; Gln_synth.
 DR Pfam; PF00120; gln-synt. 2.
 DR Prodom; PD001057; GlnA_adenyltn. 1.
 KW Ligase; Plasmid.
 SO SEQUENCE 492 AA; 54538 MW; 721A97FEA4D9CCF CRC64;

P9529 Length: 492 May 30, 2002 09:32 Type: P Check: 7071 ..

1 MSGKRIEKGK IWSDTQKAA ADVLNKIEKA GLOWRLSP DOYGLNRKM
 51 LSVAAIRSAF AGSEITMAP EFPDTASAIY ENPESADGL GSABLSPN
 101 VVMVDPPTTF RILPWADRTG WMLADLYMTS GRPALSPPA ILKRALYEMQ
 151 DLGYDYQAGL EWEYVLTIRV DPLCEPETLG GGTGAAPRK VMPAKGYSY
 201 LIENHDEVE PIMAEVROHL LALGMLRST EDEMAPSOME TTEDVMEGLD
 251 VADTVLFRN AVKQVCRRG YLASFCKPA IGFPLASGM LHOGLTARDS
 301 GANAFIPQPG EALSALGRSY VGGLEHACA ASSFTPTTIN GYRRRRPYSL
 351 APDRVTMAKD NRAAMARVIS APGDPASRYE NRIGEPANP YLYLASQVFS
 401 GIDGIRROLD PGPIQETPYA GDVTILPHNL SEALEVLETS KFFREAFGEE
 451 FIRWMHLR SEMKRFVDAE GQVDFSGDPV TWMEHREYFE LP

11AA_SEQUENCE 1.0
 ID 044310 PRELIMINARY; PRT; 59 AA.
 AC 044310;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE RNA POLYMERASE SIGMA FACTOR (SIGA).
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PCC 7120;
 RX MEDLINE=91193199; PubMed=1901566;
 RA Brahmasha B., Haselkorn R.;
 RT "Isolation and characterization of the gene encoding the principal
 RT sigma factor of the vegetative cell RNA polymerase from the
 RT Cyanobacterium Anabaena sp. strain PCC 7120.";
 RL J. Bacteriol. 173:2442-2450(1991).
 DR EMBL; M60046; AAA22044.1; -.
 SO SEQUENCE 59 AA; 6689 MW; DB46436CD0CE208D CRC64;

044310 Length: 59 May 30, 2002 09:32 Type: P Check: 4158 ..

1 MDTTRKISNS VVEDRNSWRW GTFPQAEIWN GRAMIGFLA ATLIEIFSGO

51 GFLHFWGIL

11AA_SEQUENCE 1.0
 ID 052209 PRELIMINARY; PRT; 324 AA.
 AC 052209;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PUPE PROTEIN.
 GN PUPE.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LOC5358;
 RX MEDLINE=94298771; PubMed=8026465;
 RX Koster M., Van Klompenburg W., Bilter W., Welsbeek P.;

RT "role for the outer membrane ferric-siderophore receptor PubB in
 signal transduction across the bacterial cell envelope.";
 RL EMBL J. 13:2805-2813(1994).
 DR EMBL; X77918; CAA54871.1; -
 SO SEQUENCE 324 AA; 35846 MW; EFPD238CF87C55C CRC64;
 Q52209 Length: 324 May 30, 2002 09:32 Type: P Check: 5380 ..
 1 MNGGATSIPI GEVAEQAMHW HLEIQEPAVS AATLAQMSW ROAHPLHEHA
 51 WQRTQVFAOR LREMSPGOR PLAAHALRPQ QSRRTALKQL SLMAAGAGA
 101 WYLKDALYQ DMRADYHSR GEORRLTLAD GTQYQNTDS ALVAPEQQA
 151 RRLRVNGEM LITRPALADS RPLWVDEHG RLESTLAQFN VRLHGHTQA
 201 TVYQGSVALQ PALHAYRPL IGAGQASFN QOGLIARQV AAVAPANSQ
 251 MVAQGPRLA AFIEDLARVR RGHLCADPAL AGLAVSGTFP LEWTDKIIAA
 301 VAEFLQLEVO HETRYVTLK PRMA
 IIAA SEQUENCE 1.0
 ID Q9F0G8 PRELIMINARY; PRT; 364 AA.
 AC Q9F0G8;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE ENDOLUCINASE PRECURSOR.
 GN ENDS.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mchaud P., Belaich A., Courtois B., Courtois J.;
 RT "Cloning sequencing and overexpression of a sinorhizobium meliloti
 endoglucanase gene";
 RL EMBL; AF233448; AAC44364.1; -
 DR EMBL; AF233448; AAC44364.1; -
 DR InterPro: IPR001547; Glyco_hydro_F5.
 DR Pfam: PF00150; cellulase; 1.
 DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
 KW signal.
 FT SIGNAL. 1 27 POTENTIAL.
 FT CHAIN 28 364 ENDOLUCINASE.
 SO SEQUENCE 364 AA; 39832 MW; C927C770AAE832AF CRC64;
 Q9F0G8 Length: 364 May 30, 2002 09:32 Type: P Check: 6799 ..
 1 MKSMRTTAR KGMILPLVRA CGLALTLIAT TSOAVLASGT CLRGINLAGA
 51 EFGEDGVFG TAYIYPSDET IRYFADKGFN SVRLPFMSR LQTSLNADFD
 101 AAEERLKDT VRRLREAGQI VILDPHNYAR YRGELIGSEA VYVEAFADFW
 151 SKLSLAFGNO DGIAGFLMNE PHTMPTEQWL TGANVAIAAI RSTGARNLIL
 201 VPGNSMGAI SMWGEDYGA NGVVMAGYKD PLDHYAFEVH QYLDVFSGT
 251 KDNCSTRADA IAAIENTYQW LRDNGKRGYL GQFVPPKDEA CYVALSRMD
 301 TVERGRDWTI GWAYWAGDM WPAEALNIQ PTDAGDRPOL QGLSRALSDP
 351 SPQAMTCPSL SGCG
 IIAA SEQUENCE 1.0
 ID Q06954 PRELIMINARY; PRT; 460 AA.
 AC Q06954;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

DE REBG PROTEIN.
 GN REBG.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-017 / BIOTYPE EL TOR / SEROTYPE OGAWA;
 RA MEDLINE=92212870; PubMed=1372980;
 RX Stroehner U.H., Karageorgos L.E., Morona R., Manning P.A.;
 RT "Serotype conversion in *Vibrio cholerae* O1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2566-2570(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-017;
 RA MEDLINE=95309704; PubMed=7540582;
 RX Manning P.A., Stroehner U.H., Karageorgos L.E., Morona R.;
 RT "Putative O-antigen transport genes within the rfb region of *Vibrio*
 cholerae O1 are homologous to those for capsule transport.";
 RL Gene 158:1-7(1995).
 DR EMBL; X59554; CAA42138.1; -
 DR InterPro: IPR002047; AKH.
 DR PROSITE: PS00256; AKH; UNKNOWN_1.
 SO SEQUENCE 460 AA; 52813 MW; A1AA149D3C014937 CRC64;
 Q06954 Length: 460 May 30, 2002 09:32 Type: P Check: 8871 ..
 1 MTEKKIIVW VPLPPSSWR GEGIAQTEN IVRNISPERK IEIVSSKHA
 51 EMLVGLKSN PNISVLTIGF RGSRTKKTIG YVSLNEVKD SLMDVYAKL
 101 PIIPAIFRKV GMVYSOLEYL LSLYIYSHQ RRGRESSNMC RVMLPPIIP
 151 YTHLLGGEKE VSEWDPFVFE YNKEPPLPAE YFVKLKHFE SNASAITQS
 201 RANKDYLETV MGTESSKINV IYNGSPDYSE FKQOQNSNF SEWMSKEFS
 251 GASKRAAFEA LVNHQLNFSV LWRLLTKNV SNRKIVLIST QNRPYKGFQ
 301 LEVLINICL RRDNYDFIFT CWPPTKLKER YPSYERIEH VTRVDYTLHA
 351 SLVMSDIIVL HPSNVGEGG AVPOYEASSV GRPSLINGR HVNEMEEGF
 401 DVIDLSSNFV NREYVDKIE KLINSEBYMR QNIDAINRLK ISMKESASNY
 451 ENVEFGNENA
 IIAA SEQUENCE 1.0
 ID Q00620 PRELIMINARY; PRT; 683 AA.
 AC Q00620;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE OUTER MEMBRANE PROTEIN NOSA PRECURSOR.
 GN NOSA.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91358317; PubMed=1885521;
 RA Lee H.S., Abdelal A.H., Clark M.A., Ingraham J.L.;
 RT "Molecular characterization of nosA, a Pseudomonas stutzeri gene
 encoding an outer membrane protein required to make copper-containing
 N2O reductase.";
 RL J. Bacteriol. 173:5406-5413(1991).
 CC -FUNCTION: REQUIRED TO MAKE COPPER-CONTAINING N-2-O REDUCTASE.
 CC -SUBCELLULAR LOCATION: OUTER MEMBRANE.
 DR EMBL; M60717; AAA25906.1; -
 DR HSSP; P05825; IPEP.
 DR InterPro: IPR001005; Myb_DNA_bind.
 DR InterPro: IPR000531; TonB_boxC.

DR Pfam; PF00593; TonB_boxC; 1.
 DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
 KW Signal; Outer membrane.
 FT SIGNAL 1 44
 FT CHAIN 45 683 OUTER MEMBRANE PROTEIN NOSA.
 SQ SEQUENCE 683 AA; 74924 MW; 75652273ADC4CF1 CRC64;

000620 Length: 683 May 30, 2002 09:32 Type: P Check: 5873 ..

1 MSSTRGCGA RARLSERFSI QPSRLRMQIA HATILGLLAS GALAAESVDH
 51 SEAHHAHSSA ELAPWITGV AQSPLTVAT DPKIPQVPV ASDAGDYLQT
 101 IPGSAVARGG GSNSDPYFRG MEGSRILKLLA NGAEMLGACP SRMDSPESSI
 151 TPENYDALTV IKGPQTVLMG PGNASATILL ERDPEDFSEL GGRIDASFIV
 201 GSDGRFDRIIT DAAAGGEQGY IRLANRSDS DDYDGDGDD VHSWDXKWT
 251 DLVIGWTPDE DTLLLETVGR GCGEARYAGR MWDGQSFERE SVALRFEKTN
 301 LGEHLKRIEA RYYNYADHY MONTSLRTPP MMQATNTDR RTLGGRMAAT
 351 WQLEDEYELVT GVDAGTNEHR RRGVDYKSK PMEKDADPHN YGLEFGEIIRT
 401 LNDSDRYVIG ARLDHATKAD YRSTGPSAGD SRSDNLPSCF LRYEHDQLSL
 451 PATAVVGIGH TORFPDYWEL FSGADAFER LDPEKTQLD FGLQYSKPL
 501 DAWYSAVYQG VRDYILFYSY PSKYSENIDA RIMGELCAT YRLTSNNKTD
 551 ASLAVANGKN SSDGALPOM PPLEGRIGLT YEQGDMSAAG LMRVVAQNR
 601 VAEGKGVNTS KDPDESGFG VESLNGAYRV NQNFKLSTGI DNLEDKAYSE
 651 HLNQAGNAGI GLSADERINE PGRTMARVD MSF

11AA_SEQUENCE 1.0
 ID Q930E9 PRELIMINARY; PRT; 567 AA.
 AC Q930E9;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE PHB SYNTHASE.
 GN PHB.
 OS Azotobacter vinelandii.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Azotobacter.
 OX NCBI_TaxID=354;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DW 136;
 RX MEDLINE-20578894; PubMed-11137297;
 RA Segura D., Vargas E., Espin G.;
 RT "Beta-ketothiolase genes in Azotobacter vinelandii.";
 RL Gene 260:113-120(2000).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-DW 136;
 RA Segura D., Espin G.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF267243; AAK72597.1; -
 SQ SEQUENCE 567 AA; 64404 MW; 6A02ABC102EC9482 CRC64;

0930E9 Length: 567 May 30, 2002 09:32 Type: P Check: 4054 ..

1 MDQAPSTSF WSAQTPPEVAN PALQQLRLWV NTSPWFVDTD NAKWFDIPAE
 51 RLHIQSDYQ REMYQLGIQL LSRKPSFTD KRFASDNSS PLFESLAAVY
 101 LLNSRYMML LELKIEDEK PRORLYLYVE QAIASASNN FFASNPDALE
 151 ALVKSNGVSL FNGMLHLASD LKEGKLROCD SGDPQVGRDV ATPGDIIVE

201 NEFQLIQYR PLSEKQYORP LLIVPESINK YYILDRPEN SLVRYALEOG
 251 HOVELYSWRN FDASCAGKTW DNFIDQAAIK AIKYTRAIASG GQPLNCGFC
 301 IGGTLLSTAL AVLEAGSKD HISSLTLLAT FLIDYSDPGVI NVEVDEQFVT
 351 QRERTIGCG GYVGLFRGOD MCNTFSLLR NELMANNYVD KYLAGOKPRT
 401 LDMLEFNND S TNLGPRYCW YLRHTYLQND LKSGGLECCG VRDLSTIKA
 451 PYYLGTODD HIYWRSAVN TSNLSGKIR EYLGASGHA GVINPAOK
 501 RHYWTNEQTP ADPDIMMETA EKKPGSMWD WFAWLVOHAG EQGPAYKSG
 551 NREYVIEAA PGRYVK

11AA_SEQUENCE 1.0
 ID Q93N64 PRELIMINARY; PRT; 314 AA.
 AC Q93N64;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE HYPOTHETICAL 36.1 KDA PROTEIN.
 OS Coxiella burnetii.
 OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
 OC Coxiella group; Coxiella.
 OX NCBI_TaxID=777;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Hoover T.A., Vodkin M.H., Williams J.C., Culp D.W., Thompson H.A.;
 RT "A chromosomal DNA deletion explains the phenotype of the Coxiella burnetii phase II variant.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF387640; AAK71258.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 314 AA; 36137 MW; EAE20BCDA772D2A7 CRC64;

093N64 Length: 314 May 30, 2002 09:32 Type: P Check: 559 ..

1 MKTLYVGIGV QGKKRRFPAG NDYGVSDPI NHEADYAGIH EVPLSDYEAA
 51 LVCLPDEPKY EVIRYLLSHD KHVLEKPLM TPERKQFEL QKLAQEHNVV
 101 LYTANNHRE PHIKISELL KENVLGEYVR CLFLYNGNGA KLVNDSGMRD
 151 AGGVLTLDC CHLIDFVDYW FGLRDSNYVC IDACCHENNA PDHVIILNRV
 201 AKIKIEFEMS LLSWRNSFYC DIIGEGSLH IDSLCKWGPS QLFPRORALP
 251 SGRPEKETT LIQSDPTWEL EYVYFKSLIE AKQATNLEKD IWITDQNLGL
 301 TQALKEIDL CTVO

11AA_SEQUENCE 1.0
 ID Q9P981 PRELIMINARY; PRT; 385 AA.
 AC Q9P981;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE ENDOGLUCANASE 1.
 GN EG 1.
 OS Robiliarda sp. (strain Y-20).
 OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Robiliarda.
 OX NCBI_TaxID=72589;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-Y-20;
 RA Kashiwagi Y.;
 RT "Endoglucanase gene from cellulytic fungi, Robiliarda sp. Y-20.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB030819; BAA90480.1; -
 DR InterPro; IPR001547; Glyco_hydro_F5.

DR Pfam: PF00150; cellulase.1.
 DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
 SQ SEQUENCE 385 AA: 41428 MW: 149604D42369ADJ3 CRC64;
 Q9P981 Length: 385 May 30, 2002 09:32 Type: P Check: 4445 ..

1 MKLVFSLAS LLSGASATIV YAGVAESSGE FGWMSAQTPT GTGLPRFGV
 51 DYAFISEAAV DVHVDONHIN LFRVAFLLER MCPPATSLGA AFENETHDFY
 101 KEAVDYITTV KGAVALIDPH NYMRYNDPSY QPFSGSYIGN TSDSTAATTE
 151 QPFGEWGLA SRFNDNERVI FGLMNEPHDM ATSLVLANNQ AALDAIRAN
 201 ASNLIIPGN SWTGHSWTE GSDPSALLN QFKDPLNNTA IDIHEYLDD
 251 FSGGHLECVS DEPTNLALT AMLKENNLKA FITEFGGSNS TSCQMLPDL
 301 INYMAONAEY IGWTAMAAGR FMGRNPSPCT NSTQLGSLER GSTAVDGSFG
 351 LYDTWLPVI QPLVPTLELW SGPASISGGE LTSRA

11AA_SEQUENCE 1.0
 ID Q9UW21 PRELIMINARY; PRT; 429 AA.
 AC Q9UW21;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE HYPOTHETICAL 49.7 KDA PROTEIN.
 GN OBPALEPA.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxId=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SC5314;
 RX MEDLINE=99387068; PubMed=10455055;
 RA Hull C.M., Johnson A.D.;
 RT "Identification of a mating type-like locus in the asexual pathogenic
 yeast Candida albicans."
 RL Science 285:1271-1275(1999).
 DR EMBL: AF167163; AAD51409.1; -;
 DR InterPro: IPR000648; Oxysterol_BP.
 DR Pfam: PF01237; Oxysterol_BP; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 429 AA: 49741 MW: 259673A904D5795D CRC64;
 Q9UW21 Length: 429 May 30, 2002 09:32 Type: P Check: 7245 ..

1 MGLTAKLDK KUKTSEDE DDSGAKPDN SDDIDEVDSE YONILLSTIA
 51 QLRPGMDLSK ILLPFILEK KSMLERTNF FOIPIKLIDS NSIEDPLDRF
 101 IGVLRWYLA WHISPAVAK PLNPVIGEV TCWDELPHN KSAVYLSEQL
 151 SHHPSSSYF YIMPEKIRV DGVIYKSRF LGNSSAIME GCGYVTLGN
 201 DNEVVMNOR NYVIGILFG KMRTELGDH YKRCERNGLE ANIEFKTGF
 251 IYGTDAIEG IIKDSETOKE LFOISGKWE VMYIKNIKTG KKEVALDTG
 301 SKTLAKVRP LEEQWDESR KLMKPTINGL AKRNHEALATE ESKSVENEDR
 351 IKAKKLEDD VEFHKEFERE VNENDNGYKN LEVVIYKFFD LKEDPEVLR
 401 RLFTVAPIV GQKPEKPHY PAFKKPESN

11AA_SEQUENCE 1.0
 ID Q9U123 PRELIMINARY; PRT; 190 AA.
 AC Q9U123;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE PRO0529.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Zhang C., Yu Y., Zhang S., Ouyang S., Luo L., Wei H., Zhou G.,
 RA Zhang Y., Liu M., He F.;
 RT "Functional prediction of the coding sequences of 9 new genes deduced
 by analysis of cDNA clones from human fetal liver."
 RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF11848; AAF16687.1; -;
 DR SEQUENCE 190 AA: 21480 MW: 4B8104A29AA3844 CRC64;
 Q9U123 Length: 190 May 30, 2002 09:32 Type: P Check: 6679 ..

1 MCLCVCCLYVC ICYYVCVCHP VCFWVCLPLR VSVYLYLRVC VCVCVFVCLC
 51 MCVRGCVSYVC VCVCIEREGE RKGATDGSAM KYVPHSQPWE ESNVPTGQD
 101 QLWMCCLADSG NTFHLRMGL HFLGKECRSM SLKCEFFPF VIRAPCVH
 151 WLTVNLRVG DSHRETEGT ADSEQESGCT SLPLGPNPOL

11AA_SEQUENCE 1.0
 ID Q9NZH3 PRELIMINARY; PRT; 280 AA.
 AC Q9NZH3;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE C-TYPE LECTIN-LIKE RECEPTOR-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20135876; PubMed=10671229;
 RX Colonna M., Samaridis J., Angman L.;
 RT "Molecular characterization of two novel C-type lectin-like receptors,
 one of which is selectively expressed in human dendritic cells."
 RL Eur. J. Immunol. 30:697-704(2000).
 DR EMBL: AF200949; AAF36830.1; -;
 DR InterPro: IPR001304; Lectin_c.
 DR Pfam: PF00059; Lectin_c; 1.
 DR SMART: SM00034; CLECT_1.
 DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 DR Receptor; Lectin.
 KW Receptor; Lectin.
 SQ SEQUENCE 280 AA: 32039 MW: 123C2C32FD5F8F6F CRC64;
 Q9NZH3 Length: 280 May 30, 2002 09:32 Type: P Check: 2970 ..

1 MQAKYSTRD MDDDGDTTM SLHSOASATT RHPRPRTTEH RAPSSTWBPV
 51 ALTLTLCLV LLIGLAAMGL LFFOYQLSN TGODTISOME ERLGNTSOEL
 101 QSLQVONIKL AGSLQVAKK LQRELYNKAG AHRSPCTEQ WKHGNOCIO
 151 FYDSKSWED KYFCLSENS TMLKINKOED LEPAASOSYS EFFEYSWTGL
 201 LRPDGKAWL WMDGTPFTSE LFHIIIDYTS PRSHCVAIL NGMIFSKDCK
 251 ELKRCVCERR AGMYKPESLH VPPEILGSD

11AA_SEQUENCE 1.0
 ID Q9H245 PRELIMINARY; PRT; 531 AA.
 AC Q9H245;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE C10RF28.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21218927; PubMed=11318611;
 RA Sood R., Bonner T.I., Malakowaka I., Stephan D.A., Robbins C.M.,
 RA Connors T.D., Morgenbesser S.D., Su K., Faruqi M.U., Pinkett H.,
 RA Graham C., Baxevanis A.D., Klinger K.W., Landes G.M., Trent J.M.,
 RA Carpten J.D.;
 RT "Cloning and characterization of 13 novel transcripts and the human
 RT RGS8 gene from the 1q25 region encompassing the hereditary prostate
 RT cancer (hpc1) locus.";
 RL Genomics 73:211-222(2001).
 DR EMBL: AF312865; AAG45339.1; -
 SQ SEQUENCE 531 AA; 60615 MW; EBF7BAA61A5270B CRC64;

Q9H245 Length: 531 May 30, 2002 09:32 Type: P Check: 5880 ..

1 MADVSLVRQ YNCKKEIV KQDEVEGEF SMPKNKTVY VVWGGRGEGQ
 51 PREYTLDSI LFLNNVHLS HPVYVRRAP ENIPVVRPD RKDLIGYING
 101 EASTASIDR SAPLEIGLQ RSTQVKRADE VLAFAKPRI EDEECVRLDK
 151 EBLARLESH KEGIVOTEQI RSLSEAMSV KTAIAIKAKIM AKKRSTITD
 201 LDDDTALQK RSYDAEYDV TROIVSERY WRTTITLOS TGNPSKINF
 251 AIIQSVKARE EGRAEORPA PNAAPVDPTL RTKQIPPAV NRYDQERFG
 301 KETEGFKID TMGTYHGMLT KSVTEGKSAR KTGTPAAGV PRPVQARPP
 351 PNOCKGSRTP IITIPATTS LITMLNAKDL LODLKFVPSD EKKKQCGRE
 401 NETLIORRD QMOPGTAIS VTPYRVVDQ PLKMPQMD RYVAVFVQSP
 451 AMQFGWPL LRDGSPYDIF AKIKAFILKY DEVRILDPNV KMDVIVLELS
 501 YAKRLDRPV FLRFWEITDR YWVKHSHLR F
 11AA_SEQUENCE 1.0
 ID Q9BXJ8 PRELIMINARY; PRT: 343 AA.
 AC Q9BXJ8;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE TRANSMEMBRANE PROTEIN INDUCED BY TUMOR NECROSIS FACTOR ALPHA.
 GN TMPT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Murakami T., Matak C., Hamakubo T., Kodama T.;
 RT "Endothelial cell transmembrane protein induced by tumor necrosis
 RT factor alpha (TMPT) mRNA, complete cds.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF327923; AAK16442.1; -
 SQ SEQUENCE 343 AA; 40610 MW; 7AGE241804F59A6 CRC64;

Q9BXJ8 Length: 343 May 30, 2002 09:32 Type: P Check: 1303 ..

1 MOPPPGPG DCLRDWEDLQ QDFQNIQETH RLYRLKLEEL TKLQNNCTSS
 51 ITRQKRRLQ LALALRKCP SLPAEAGAA QELQNMKER QGLFDMEX
 101 LKKNGLIYS LVLGWVNTL LSKQAKFAY DEYEKKLIL TIIILLIST
 151 CFFLNSRVT DAAPFLLWV YCTLTITRES IILNNGSRIK GMPVFNHYVS

201 TELSGVMTW PDGLMYQKFR NOFLSPSMQ SFVQFLQYVY QSGCLYRLRA
 251 LGRHTMDLT VEGQSMWNR GLTFILPFL FGHFWQLFNA LTLFNLADDP
 301 QCKEWOVLNC GPFILLFLG NFFTLLRVH HKFHSORHGS KKD

11AA_SEQUENCE 1.0
 ID Q9NV66 PRELIMINARY; PRT: 732 AA.
 AC Q9NV66;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CDNA FLJ10900 FIS, CLONE NT2RP5003522, WEAKLY SIMILAR TO
 DE NADPH-CYTOCHROME P450 REDUCTASE (EC 1.6.2.4).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hirooka S., Ishii S., Kawai Y.,
 RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K.,
 RA Masuho Y., Kanehori K.;
 RT "NEO human cDNA sequencing project.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK001762; BAA91891.1; -
 DR HSP: P16435; IBC.
 DR InterPro: IPR001094; Flavdxn-1like.
 DR InterPro: IPR001226; Flavodoxin.
 DR Pfam: PF00258; Flavodoxin. 1.
 DR PRINTS: PR00369; FLAVODOXIN.
 SQ SEQUENCE 732 AA; 83720 MW; F8C92411EC2B339C CRC64;

Q9NV66 Length: 732 May 30, 2002 09:32 Type: P Check: 4753 ..

1 MDSADTDL FSPSLISWIN RFYILGFV SISLWICVQI VIKTGKNLQ
 51 EKSVPKAAOD LMTNGVSLQ EKDIFFVSGV IFGSGOTGTA KGFATVLAEA
 101 VTSLDLPAI INLKEVDPD HLIBEVTSKN VCVFLVATYV DGLPTESAEW
 151 FCKWLEASI DERFGTYLK GMRVAVFGLG NSAYASHPK VCKNDKMLM
 201 MGAHVYMSR GEGCCDVYKS KHGSTEADFR AKTKTITSL QALQGERKK
 251 SCGGHCKGR CESHQGSSE REEGSHODE LHHRDTEEEB PRESSSEEF
 301 GGEDHQSLS IYDVEDLKI MDHYKKEKE KEOQEEKSGI FNNMGRNEDG
 351 ERRAMTTPAL REALTKQGY LIGSHGVKL CRTTKSMLRG KGCYKHTFY
 401 GIESHRCMET TPSLACANKC VFCWRHNTNP VGTEWRMKMD QPEMLKEAI
 451 ENHNNIKQF KGVPGVKAER FEEGMTVKHC ALSLVGEPII YVEINRFLKL
 501 LHOCKTSSFL VTNAOPRAEI RNLEPVQLY VSYDASTKDS LKKIDRPLFK
 551 DFQREFLDSL KALAVKQORT YRRLTLVKAW NYDELQAYAO IYSLGNPFI
 601 EYKGVATGCE SSASSLTMAN VPMHEEVQF VRELVDLPIE YLACAEHNS
 651 NCLLIHRRF KIGGEWWTWI NYNRFQELIQ EYDSDSGST PSAKDYMAFT
 701 PHWALFGASE RGFDPKPTRH QRKNKSKAIS GC

11AA_SEQUENCE 1.0
 ID Q9H9C4 PRELIMINARY; PRT: 561 AA.
 AC Q9H9C4;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)
 DE CDNA FLJ12855 FIS, CLONE NT2RP2003506, WEAKLY SIMILAR TO
 DE MAPH-CYTCHROME P450 REDUCTASE (EC 1.6.2.4).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Makematsu A., Nakamura Y., Nagahari K., Masuno Y., Oshima A.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK022917; BAB14307.1;
 DR HSSP; P16435; 1B1C.
 DR InterPro: IPR001226; Flavodoxin.
 DR Pfam; PF00258; flavodoxin.1.
 SO SEQUENCE 561 AA; 64682 MW; 4DF7641582E575F8 CRC64;

09H9C4 Length: 561 May 30, 2002 09:32 Type: P Check: 6460

1 MRYAVFGLN SAVASHENKV GKNVDMKIM LGAHVMSRG EGDGDVYKSK
 51 HGSIEADPRA WKTFISLOQ ALQKGERKKS CGCHCKKGC ESHQSGSEER
 101 EEGSHEDDEL HHRDTEEEP FESSSEEEFG GEDHOSLMSI VDVEDLKIM
 151 DHVKKERKEK EQGEKSGLF RIMGREDE RRAMTPALR EALTKQGYQ
 201 IGSISGVKLC MWTKSMERG GGCYKHTFYG IESHRCMETT PSLAYANKV
 251 FCMRHHTNPV GTEWRMKMD PEMILKEALE NHQNMKQFK GVPGVKAEKF
 301 EEGMTVKHCA ISLYGEPIM PEINREFLKL HCKKISSFLV TNAQPAEIR
 351 NLEPVTQLYV SVDASTKDSL KRIDRPLFKD FMQRPDLSLK ALAVKOQRTV
 401 YRLTLVKAMN VDELOAYAOQ VSLGNPDPIE VKGYTCES SASSITMAHV
 451 PMHEVVOFV RELVDLPEY ETACEHEHNS CLLIHRKFK IGGEWMTWIN
 501 YNRQELIQE YEDSGGSKTF SAKDYMARIP HMAIFGASER GEPDKTRHQ
 551 RKNKSKAISG C
 11AA_SEQUENCE 1.0
 ID 09H5L7 PRELIMINARY; PRT; 355 AA.
 AC 09H5L7;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)
 DE CDNA: FLJ23316 FIS, CLONE HEP12031.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Ohtani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK026969; BAB15608.1;
 SO SEQUENCE 355 AA; 40699 MW; 84ABF38A6EB2E7C8 CRC64;

09H5L7 Length: 355 May 30, 2002 09:32 Type: P Check: 1905

1 MSVERIACNQ TKIMAKRST IKTDLDDIT ALKQSFYDA EVDVTRDVIS

51 RERWRIRTT ILOSTGKNFS KNIFAILQSV KAREGRAPE QRPAPNAAPV
 101 DPTLTQKQPI PAAYNRYDOE RFKGEKEFEG FKIDTMGYH GMLKSVTEG
 151 ASARTQTPRA AQPPRRVSO ARPPRQKKG STPIIIIIIPA ATTSLKTMNL
 201 AKDLQDLKF VPSDEKKGQ CORENETLIO RKKDOMQPG TAIIVTPYR
 251 VVDQPLKMP QDMDRVAVF VQPAWQFG WFLPLDPS VDIKAKIAF
 301 HLKYEVRILD PNVQKMWTV LELSYKRNH DRPVLREWE TLDRYMKKH
 351 SHLR
 11AA_SEQUENCE 1.0
 ID 096KN6 PRELIMINARY; PRT; 1449 AA.
 AC 096KN6;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
 DE PUTATIVE TRP CATION CHANNEL.
 GN KNP3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wehage E., Eistfeld J., Heiner I., Juengling E., Zilt C., Lueckhoff A.;
 RT "Splice variants of UTRPC2 differentially activated by ADP-ribose an
 RT hydrogen peroxide.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ417076; CAD01139.1;
 SO SEQUENCE 1449 AA; 165328 MW; 563DA8259F5A9E7 CRC64;

096KN6 Length: 1449 May 30, 2002 09:32 Type: P Check: 6589

1 MEPSALRKAG SEDEGEFGL PRRVTDLGMV SNLRSSNSL FKSRLQCFP
 51 GNDKQESLS SWIPENIKKK ECVYEVESK LSDAGKVVQ CGYHEQHL
 101 EATKPHFGQ TQMDPKKHQ EMPTDAFGDI VFTGSQYVK KYAVNSQDTP
 151 SSVYIHMTQ HMGIDVPLNL ISVTGAKNF NKKPRLKSF RGLVKAQT
 201 TGAMITIGS HTGYMKQVE AVPDFSLSS YKEGELITG VATGTVHRR
 251 EGLIHPGFS PAEYILEDG QGNLTCLDSN HSHFLVDDG THGYGYEIP
 301 LTRLEKFTS EQTKERGVA IKPIYCVLV EGGPETHHTI DNATYNGTPC
 351 VVVEGSGRVA DVIAOVANLP VSDITISLIQ QKLSVFQEM FETTESRIY
 401 EMTKKIDQIV RRQQLLVFR EKGDDQDQVD VAILQALLKA SRQDHEGHE
 451 NMDIOLKLV AMNRVDIARS EIPMDQWK PSDLHTMTA ALLSNKPEY
 501 KLFENGVOQ KERYVTDTLL YLEENIDPSC LFHSKLQNH VAQVRLRLG
 551 DFTQPLVPR RHNDRLRLLL PVPHVLANV GVSLSLTKR SSGVITTTMD
 601 PIRDLITAI VONRELQGI IMNOSODCIA AALACSKTLK ELSKEEDTD
 651 SSEEMLALAE EYEHRAIGV TECYRDEBR AQKLITRASE AMGTTCLQ
 701 ALEAKDMKFV SHGIGQAFI KYWVGQLSVD NGIMRVITCM LAPPLLITGL
 751 ISFREKRLQD VGTPARARA FFTAPVVFH LNIISYFAFL CLRAYVLAVD
 801 PQPVPWCCE AIYIMLFSLV CEMRQLFYD PDEGLMKKA ALYSDFWNK
 851 LDVGAILLFV AGLTCLRIPA TLYPRVVIS LDFILFCLRL MHFTTSKTL

901 GPKIIIVKRM MKDVEFFELF LAVWVVSFGV AKOAILIINE RRVDMLEFGA
 951 VYHSYLTIFG QIPGYIDGVN FNPHECSPNG TDPYKPCPE SDATQQRPAF
 1001 PEMLTVLLLC LYLFTNILL LNLILAMFNV TPQOVQEHFD QIMKFOHDL
 1051 IEEYHGRPA PPPFILLNHL QEFIKRVYVK TPARKRKQVK NLEKNEEA
 1101 LLSWEIYLKE NYLONROFOQ KORPEQKIED ISNKVDAMVD LLDLDPLKRS
 1151 GSMQORLASL EEOVAQTARA LHMIVRTTLA SGFSSEADVP TLASOKAAEE
 1201 PDAPGGGRKK TEPPGDSYHV NARHLLYPNC PVTRPPVNE KVPWTEFLI
 1251 YDPFYTAEK KDAAMPDPMG ENPMGRTGLR GRGSLSCFGP NHTLYPVTR
 1301 WRNEDDGAIC KRSIKKMLEV LVYKPLSEH WALPGSREP GEMLPRLKLR
 1351 ILROEHMPSF ENLLKCGMEV YKGYMDPPRN TDNAWETIVA VSVHFDOND
 1401 VELNRUNSNL HACDSGASIR WQVVDRIPL YANHKTLLQK AAAREFGAHY

11AA_SEQUENCE 1.0
 ID Q96JM4 PRELIMINARY; PRT: 1227 AA.
 AC Q96JM4;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE KIAA1801 PROTEIN (FRAGMENT).
 GN KIAA1801.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE-21245130; PubMed-11347906;
 RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XX.
 RT for large Proteins in vitro."
 RL DNA Res. 8:85-95(2001).
 DR EMBL; AB058704; BAB47430.1; -.
 FT NON_TER 1
 SO SEQUENCE 1227 AA; 142754 MW; A20651382DC61B8F CRC64;

Q96JM4 Length: 1227 May 30, 2002 09:32 Type: P Check: 5997 ..
 1 RTREFDOQEK EEFMRSTKTDG ATPDFVPEPS PHDLPMDEHV LPDDADINFG
 51 YCEVEBCKRQ SFEMQEKOK ELEDKEKQTL KAQRNREBKQ FOEEBKRHC
 101 WMKQFYVEKK KLENIOKQEQ DKMDELKYE EKIWEKEFKQ HEEYIRNLHL
 151 QMEERTFRK DQOEKKNL LKQONNAAYK IQAKYKAFVA YQYKGPITKE
 201 QIESKRRKQ EMKEKEKAKR QKEENRRL EEOQRIKEER KKOEERKR
 251 REKEYEKKK IVKQEREQLI SKEKILRED ASQOLIISA LKSGYNNKH
 301 LSLDISNDK GDIAKNLYDE NSKKOEDVLV WLVEESNMKE NYDRQTLIKE
 351 SIQVLMKESI SSQTLADFK MEKKNNENLAK KRCSEELVKQ ERKENENDNK
 401 TELGNSDLKG NLKEQFPLOE LKSDAKEEK IMKHVINENT GQKTOIILGH
 451 NOEISEVKTN EEOKIIRKQO OKKIQEVEK EIQEONGILY KDKDTLIVSY
 501 KORSLSLSE NSKDVRENVI LOEKELYSKS KEIEENPRDN AMNGCIYFN
 551 TTDIMINIEG KRNDQYVLG RHAPCEGLSN YNAESSMYSK EVNSLAKSEIR

601 NISEKCHENA PEPSMTQCV SESTLLYSIE ERLIMIXSF KPWLEIFKON
 651 QOKKIVRRKR PVKCPAMTP ALDKLEILRG GPMWDLQVY TVTQDILRG
 701 VLSLTAECTN LQFLSLRRCG LLSLSLSNC KKLKIDAOE NHIDALCEN
 751 LENICVVLN KNOTLSLGH DGCTNIOCLE LSYNKITRIG YSFLIEKLY
 801 DNTGCHHLG TSSTYSLSAQ VWIPYGLCWS WIPITSILKN SDCNFLSHL
 851 YWNCLESK NLQOLLDHN QLINTKGLCD TPTIYLLDS HNHLTDYEGV
 901 ENCGILQILK LOGNYLSELP SLENVLBRE LHLDNSIST VEAFSYWLP
 951 LQNTITSON SLTKIYPLFH FVSLKLDVS HNCISDLKSA IKWPDACYSL
 1001 HELSLTGMP LQETNMRDNL LKYLPAIRIL NGNLIINSSE SRPEHNQIG
 1051 SAGFLALCOS QIREFNLLIE NYITGKGVF TLDTAENLCH YFKKILMIST
 1101 EYRHAHERGD VTIYKDESE AOKNHLAPTN SDSTIQNGVF YSCAREGEPD
 1151 SPDIPEKAMD SVSSHSPLSK SATCENMEGR HOELIVCQR EDSKASSIPT
 1201 IRIPEKEYVM TNSILRNHON IEPSEKM

11AA_SEQUENCE 1.0
 ID Q96A07 PRELIMINARY; PRT: 355 AA.
 AC Q96A07;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE SIMILAR TO HYPOTHETICAL PROTEIN FLJ23316.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLADDER CARCINOMA;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RC TISSUE-LUNG CARCINOMA;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC014351; AAH14351.1; -.
 DR EMBL; BC011808; AAH11808.1; -.
 KW Hypothetical protein.
 SO SEQUENCE 355 AA; 40621 MW; 3F228230AB6F3E9F CRC64;

Q96A07 Length: 355 May 30, 2002 09:32 Type: P Check: 1525 ..
 1 MSVERIAIAK AKIMAKRST IKTDLDDIT ALKQSFYDA EVDYTRDVIS
 51 RERWARTRTT ILOSTGKNFS KNIFALQSV KAREGRAPF QRPANAPV
 101 DPTLTQKPI PAAYNRYDQF REFKEETEGF FKIDTGTYH GWTLSKYTEG
 151 ASARKTQTPA AQPYRPVSO ARPPNQKGK SRTPIIIPA ATSLITMLN
 201 AKDLQDLKF VPSDEKKKG CORENTLIG RKDDOMQCG TAISVTPYR
 251 VVDOPLKIMP QDMDRVAVF VQGPANQFG WPMLLPDGSP VDIKAKAF
 301 HLKXDEVRLD PNYQKMDVTY LELSYHKRHL DRPYFLRWE TLDXYAVNKH
 -351 SHLR

11AA_SEQUENCE 1.0
 ID Q9V696 PRELIMINARY; PRT: 2040 AA.
 AC Q9V696;

DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE CG8487 PROTEIN.
GN CG8487.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Tracheata; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sultion G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agdayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosløe C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,
RA Palatzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Sjöler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003822; AAF58532.1; -;
DR HSSP: Q99418; 1PBV.
DR FLYBase: FBgn003714; CG8487.
DR InterPro: IPR000904; Sec7.
DR Pfam: PF01369; Sec7; 1.
DR SMART: SM00222; Sec7; 1.
DR PROSITE: PS50190; Sec7; 1.
SQ SEQUENCE 2040 AA; 226981 MW; A59107B06813FD7C CRC64;
09v696 Length: 2040 May 30, 2002 09:32 Type: P Check: 9003 ..

1 MALPGNGIY VNGEMATLMT AMRGRTRWNA TAYVDDENDS LKLFLDLKH
51 ELNRIEDLRQ IEPQVLAPE LEVIRTDAT GPLTSLALAS VNKLSYGLI
101 DPTSPNLADI VERIADAVH ARFAGTDOSS DGYTFMAYIE VLHLLRSPE
151 GAAYSNVSMC EYMLSCFKIS FEPRLSELLR RSAEKSLSKM VLLFPMRLPQ
201 FAERSDPTML QKRFTIGDAA SGATQEKLR KTVAAQTAP RKSSAVEEPP

251 QTPQSANLTV PGHLKAPILA TTPASPAGNI LDMQKITTQT PTTATSGED
301 ETVVPEPIVI QVESTSEPL LDGETGEATS TLAEPANSEY INSGVAFETQ
351 QSTHHDVTSI SPYGLPFIQE LFRPLILCN PLDQNSDSM MHGSLSLTV
401 AFEVAADNIG KYEGLELVK DILCRNLISL LSEERLSTFA ADIQLCFLE
451 ESLRGHLKFQ LEAVLKLSE IIASDNPKPT YEMRELADN LLQLMRIPGF
501 VTELYINVC DLYCTDMFES LFNLSKYTL SATNAVYSTH IISMDTLISV
551 IDSIERNCAA SKNSNNRES LPEAAPATGC SRHSRHNSGL EGIYVDSGNS
601 VAAEKEVENI ASFINASSHR LRLOGGEGV GITSEQLAKV KOKKRLLSOG
651 TERNQPEK GIOYLQEHGI LNAELDPKV ALFLRENGI DKNIGIYIS
701 KKKRVDSKIL INFVDSFDT GLRVDAALRL YLEFRLEPGE APLIFVLEH
751 FSDHMRKONQ DPPANVDAAF RLAVAILMIN MDQHNSNAKR LNPMTLEDF
801 TKNLRGLNGG EDRDQMLAQ VFNATKNERI VMPAEQIGLV RENYQMKVLL
851 RRGDTHDGHF HYVHDASYDV EIFNTVWGS LSALSFMEDK STETGYQRTL
901 AGFSKSAIS AHYNLHSDP ALVLTICKET TLSSVEQHE PABANNETQ
951 AVNGLNGKA QAAMRVFLL VHDYDQCLRE SMKHLIDYLL QLFRKLPLK
1001 SLIVEPFCE ANKANMILLE KREKQESGL FSSLSYFSS EGREPEPYEE
1051 QDFIKIGRKC IKECQDQML QSKFVQLES LOELLKCYLA LTKAPOGHS
1101 IGLVYAEQDT VFMEPLVRI VVHNRDMRP LMPAVRDQMY LLMGASG
1151 YDYLNCIV AVAKLAIYM RNEELCPYVL QSLKMLMLK PALLRISKQ
1201 ISIGIYELLK TSAONIHSEQ DMQIIFNLE CVGAGAVPPN YDDAQLPLPP
1251 NSAKSDGAI SCGEDATAVP ERGYTSDSEI TKASAAPAVS SPSAEMWIV
1301 NNKSELITA SRQSPSLS APVTVLYN CQLDHPFA LFCMDSLAF
1351 IVRSVAHTP YNEACVRCI RIFVEACRQD GIORRKLIES AAKOKSKRR
1401 SERKPGMASS ASSSNLTLLT GDPDQJNG NABQEDIAQ RYBQLSIQL
1451 DLMVTLVTR AQIFRMMAE GCTVQSAAL WSPGCPILQ GIARLMDRR
1501 REVETHAISC LQORALLVHD LQTSCTEWC SCFHQVLEPL LNELPESNA
1551 AGOLDAALLE ESRIPTATTI SKVFLQHLTT LIEIGNAFNE LMDIIDYIE
1601 RKMVGSDDL SEMOELLKN MLVMSHVRV FHNQDGSIQ ALMELTWRI
1651 GEFLPNLKEE LEHDEDSTP AVSLDYQIR FAHSTGTTDY SLSLPYGVS
1701 PKISSISASK CDLPATQK KPAKLRSKL KLLKAAYAA APOQOLPAVT
1751 ILPRQTOVSN ELVVSAPTRP AATPILGSV ESPRSITLQ PRADVLODP
1801 PSFVFAPII VPPQPPAVTD PIPSTLLPD LVNEATAAV QATTSPTHS
1851 POEAQEPASI VOOTNIVTIN NTYNSYALEV PMADETTAEQ FGQOQOQLLY
1901 QQYQOYQAO QOOLPAPASD PAIWPISHL LAGNAYSPL KMQQASIVHS
1951 FAPYTESQAA TSGAGTAAD IYOEYVQNPY NLTLQOHPOQ QLHQOQOQOQ
2001 QOATGMANAF PAVATPANYF NVNVDPSISP PGSLELVGQO

11AA_SEQUENCE 1.0

ID 09V9K3 PRELIMINARY; PRT; 337 AA.
AC 09V9K3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CG14593 PROTEIN.
GN CG14593.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Braachycera; Muscomorpha;
OC Ephyroidae; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe C.R., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Plamkoc C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslir C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegyarm C.,
RA Jatalai B., Kalush F., Kapten G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jatalai M., Kalush C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Morkulov G., Mlshina N.V., Mobery C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Palancat K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shier B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Slater E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Stryckas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003784; AAF57285.2; -;
DR FLYBase; FBgn0033058; CG14593.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PSS00262; G_PROTEIN_RECEP_F1_2; 1.
DR SPOUSTE; 337 AA; 36989 MW; E06E7260DC3249C4 CRC64;
SQ
09V9K3 Length: 337 May 30, 2002 09:32 Type: P Check: 7480 ..

1 MCRISSEFKD ISIGSVFTL TALSGERCA IVNPLRKIQT KPLTVFAVM
51 IWIALLIGM PSVLESDIKS YVFATATGM TIEVCSPPRD PERAKFNAG
101 KALVYLLPL SIIGALYIMM AKRLHMSARN MPGEQSQMS RTQARLHV
151 ARMYVAFVV FFIQFFPYH FELWYHFYPT AEEDDEDMN VLKRVGCTS
201 FLNSCVNPVA LYCVSGVFR HENRYLCIC VKRPHLRQH STATGMNDT

251 SVMSMRSTY VGGTAGNLRA SLHRNSHGV GGAGGVGG VSGRGVSFH
301 RQDSMPLOHG NAGCGAGGG SSGLAGGRT AAVSEKR
11AA_SEQUENCE 1.0
ID 09GN87 PRELIMINARY; PRT; 209 AA.
AC 09GN87;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GLUTATHIONE TRANSFERASE GST1-1 (EC 2.5.1.18).
GN GST1.
OS Anopheles dirus, and
OS Anopheles dirus B.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
OC Anopheles.
OX NCBI_Taxid=7168, 123217;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=A.dirus;
RA Ketterman A.J., Prommeenate P., Boonchaay C., Chanama U.,
RA Leetachewa S., Prommet N., Prapanthadara L.-a.;
RT "Single Amino Acid Changes Outside The Active Site Significantly
RT Affect Activity of Glutathione S-Transferases.";
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=A.dirus B;
RX PubMed=1102837;
RA Pongjaroekit S., Jirajaroenrat K., Boonchaay C., Chanama U.,
RA Leetachewa S., Prapanthadara L., Ketterman A.J.;
RT "Genomic organization and putative promoters of highly conserved
RT glutathione S-transferases originating by alternative splicing in
RT Anopheles dirus.";
RL Insect Biochem. Mol. Biol. 31:75-85(2001).
DR EMBL; AF273041; AAC38507.1; -;
DR EMBL; AF251478; AAC17625.1; -;
DR HSSP; P30712; ILJR.
DR InterPro; IPR004046; GST_C.
DR InterPro; IPR004045; GST_N.
KW Transferrase.
SQ
09GN87 Length: 209 May 30, 2002 09:32 Type: P Check: 9284 ..

1 MDTYLLPESA PCRAVQNTAA ANGVELNKL TNLMAGEHMK PELKLPQH
51 CIPFLDNGF SLWESRAIQI YLEKYGKDD KLYPRDPQKR AVVNQRLEFD
101 MGLYQRNGD YWYPOIFAKQ PANAENKKM KEAVFLWTF LEGGYAAGS
151 DLTADLSLA ASATTEYVAG FDEAPYPNVA AMLARCKANA PGYALNAGA
201 DEFAKEMS
11AA_SEQUENCE 1.0
ID 09NL42 PRELIMINARY; PRT; 491 AA.
AC 09NL42;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BRACHYURY.
GN CBRACHYURY.
OS Ciona savignyi.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Clonidae; Ciona.
OX NCBI_Taxid=51511;
RN [1]
RP SEQUENCE FROM N.A.
RA Imai K., Takada N., Satoh N., Satou Y.;
RT "An essential role of beta-catenin in the endoderm specification of

```

RT ascidian embryo."
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB031545; BAA92187.1; -.
DR HSSP; P24781; 1XBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00423; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS50252; TBOX_3; 1.
SO SEQUENCE 491 AA; 54673 MW; 62E83F723AA03025 CRC64;

Q9BLA2 Length: 491 May 30, 2002 09:32 Type: P Check: 320 ..

1 MTSTLESCLM TSGNTNESSC VKMSLVNDNL WSRFHAIVNE MIVTKNGRRM
51 FPLVLTITIS LDPTANYSVM LDFAVAVDNNR WKYVNGEWMV GCKREPHYVS
101 CAVIHDPSPN FGSIMWKQPV GFSRYKLTKK ATGNAQOIML NSLHKTEPRI
151 HMKVGGAES QOIVATHSFA ETRFLAVTVAY QNEDVTSLSKI KYNPFAKAPL
201 DAKESRTEFE TYKDTFANAG TSQTYARVSS WTEGVNQTNG DNLPSYVHCD
251 PATRNDEGVN PMFAPINCPN PYQPARERR SSSQSRSHYN YHPSYRDAAH
301 HPMYHODIS TSIDVNAFNS HETLCYPTGY QVMTADOMPH TAIPIHDVAE
351 QGTFSLDQVS VEDMASLAFE SATSSASLIQ SDFSTVDSFD SSSFGAHWSY
401 STNHIGYNS NSPISFEFNP AHTAHYITNP YEMSPQCLN VAATENTTIT
451 ASLYHSPSP SAEIGSSPL ADAYDPTKLT SMTPLTPPS L

!!A:SEQUENCE 1.0
ID Q9BLA5; PRELIMINARY; PRT; 308 AA.
AC Q9BLA5;
DT 01-JUN-2001 (Tremblrel. 17; Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DE BRACHYURY PROTEIN (T) (FRAGMENT).
GN OILBRA-0.
OS Olkopleura longicauda.
OC Eukaryota; Metazoa; Chordata; Urochordata; Appendicularia;
OC Olkopleuridae; Olkopleura.
OX NCBI_TaxID=107032;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21347418; PubMed-11455437;
RA Nishino A., Satou Y., Morisawa M., Satoh N.;
RT "Brachyury (T) gene expression and notochord development in Olkopleura
longicauda (Appendicularia, urochordata).";
RL Dev. Genes Evol. 211:219-231(2001).
DR EMBL; AB052730; BAB33364.1; -.
DR HSSP; P24781; 1XBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS50252; TBOX_3; 1.
DR NON_TER 308 308
SO SEQUENCE 308 AA; 35259 MW; 7FD78F87C6FBA86 CRC64;

Q9BLF5 Length: 308 May 30, 2002 09:32 Type: P Check: 1820 ..

1 MNGTSDIKIE DHNHKKONI KYRLEDELM KRFSSLTNEH IYKNGRRMF
51 PVKRVYGL DENAMYSVLL DFSADNHRM KYVNGEWPVG GKPEQAPPC
101 VYHHPDSPNF GSHMKQSIG FSKYKLTNKL NNHNGSQIML HSLHKTEPAT

```

```

151 HIIVGGQOG TQILKTQTF PTKEIAVTA YONEITSLSK IKHNPFAKAF
201 LDKQREID QNNQNMNYOV PIENNPTPTQ SRKASRNKQO CSQPYEIKR
251 NONHSSNHLV VSLPQSNQIE PTYPAIYEPY SPQMDPMWPS GAGTPLLASI
301 GQMLDNWNS

!!A:SEQUENCE 1.0
ID Q9BLF4; PRELIMINARY; PRT; 403 AA.
AC Q9BLF4;
DT 01-JUN-2001 (Tremblrel. 17; Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE BRACHYURY PROTEIN.
GN OILBRA.
OS Olkopleura longicauda.
OC Eukaryota; Metazoa; Chordata; Urochordata; Appendicularia;
OC Olkopleuridae; Olkopleura.
OX NCBI_TaxID=107032;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21347418; PubMed-11455437;
RA Nishino A., Satou Y., Morisawa M., Satoh N.;
RT "Brachyury (T) gene expression and notochord development in Olkopleura
longicauda (Appendicularia, urochordata).";
RL Dev. Genes Evol. 211:219-231(2001).
DR EMBL; AB052731; BAB33365.1; -.
DR HSSP; P24781; 1XBR.
DR InterPro; IPR001699; T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SM00425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS50252; TBOX_3; 1.
DR SEQUENCE 403 AA; 45610 MW; 86A545378669B6EE CRC64;

Q9BLF4 Length: 403 May 30, 2002 09:32 Type: P Check: 7868 ..

1 MNGTQDIEKIE DHNHKKONI KYRLEDELM KRFSSLTNEH IYKNGRRMF
51 PVKRVYGL DENAMYSVLL DFSADNHRM KYVNGEWPVG GKPEQAPPC
101 VYHHPDSPNF GSHMKQSIG FSKYKLTNKL NNHNGSQIML HSLHKTEPAT
151 HIIVGGQOG TQILKTQTF PTKEIAVTA YONEITSLSK IKHNPFAKAF
201 LDKQREID QNNQNMNYOV PIENNPTPTQ SRKASRNKQO CSQPYEIKR
251 NONHSSNHLV LPTPTQIEQ NPYPALYEPY SPQMDPMWPS GAGTPLLASI
301 GQMLDNWNS GYVESHDNNS VOLNCSPPSS HLSNDDLOG TMTNATSGAN
351 GASSTLEPIN DFWATQYNSY NDYALNDHYS IDPTQSPDWS TSPNIALEMA
401 ANL

!!A:SEQUENCE 1.0
ID Q17041; PRELIMINARY; PRT; 491 AA.
AC Q17041;
DT 01-JAN-1998 (Tremblrel. 05; Created)
DT 01-JAN-1998 (Tremblrel. 05, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE HYPOPHYSICAL 56.5 KDA PROTEIN.
GN T15B7.16.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT Investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Pauley A., Gattung S.;
 RT "The sequence of C. elegans cosmid T15B7.";
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 DR EMBL: AF022985; AAB69969.1; -;
 DR InterPro: IPR000188; GABAA_receptor.
 DR InterPro: IPR001175; Neur_channel.
 DR Pfam: PF02931; Neur_chan_LBD; 1.
 DR Pfam: PF02932; Neur_chan_memb; 1.
 DR PRINTS: PR00252; NRIONCHANNEL.
 DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
 DR Glycoprotein: Hypothetical protein: Ionic channel;
 KM Postcytoplasmic membrane; Transmembrane.
 KW SEQUENCE 491 AA; 56494 MW; FAFB9922F24F6AEA CRC64;
 SQ

017041 Length: 491 May 30, 2002 09:32 Type: P Check: 8670 ..

1 MEVFKQMLK KLTQPSPRE HSFPEPRAND TISQVLRL FDDNYKNNTY
 51 PLTSATVSV VEEGIONIAQ VSEISASFTL DLFSQIMHD PLRFDHLIN
 101 CLQNLTLGVS MVEKLTPTNV CFVNSKTEI HSSPTPNIFL MIVPNCIVWV
 151 NYRLQVQSPC MVDLVLPMD IMNCCELLIES YAYNAAYKL NMRNQPVFS
 201 IAKSKLSDEFT LYGLOWTKNS PEYAAGQMDQ VSFCSLNCF SKMDGQQLIV
 251 SLTFSRAYGF YILQYIPTV SSVPLSPVSF WIDIKALPAR IITGVSSILMA
 301 LITQYGVAK NLPRGVYKS IDVYAVLTA FIFLTMIEVA FVGYLDSENN
 351 LRRKROAEK KKEVAVLOR KDKRRKNYGG ATVTNANAA ESDLRSNYDE
 401 PFSQNGTTSK KISISQANN MESLHALAQ FGLLTDDDE NTKWTAQNV
 451 KECKKAPPLS FCFNLITVWC YLYQNYLAK AEALAQMTTP S
 11AA-SEQUENCE 1.0
 ID Q9XYS1 PRELIMINARY; PRT; 438 AA.
 AC Q9XYS1;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE BRACHYURY PROTEIN.
 GN BRA.
 OS Clona intestinalis.
 OS Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
 OC Clonidae; Clona.
 OX NCBI_TaxID=7719;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97195701; PubMed=9043074;
 RA Corbo J.C., Levine M., Zeller R.W.;
 RT "Characterization of a notochord-specific enhancer from the Brachyury
 RT promoter region of the ascidian, Clona intestinalis.";
 RL Development 124:589-602(1997).
 RN [2]

RP SEQUENCE FROM N.A.
 RA Corbo J.C., Levine M., Zeller R.W.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF123247; AAD21079.1; -;
 DR HSSP: P24781; 1XBR.
 DR InterPro: IPR001699; T-box.
 DR Pfam: PF00907; T-box; 1.
 DR PRINTS: PR00937; TBOX.
 DR SMART: SM00425; TBOX; 1.
 DR PROSITE: PS01283; TBOX_1; 1.
 DR PROSITE: PS01284; TBOX_2; 1.
 DR PROSITE: PS50252; TBOX_3; 1.
 SQ SEQUENCE 438 AA; 49625 MW; 1DDFCB65C2D92F4F CRC64;

Q9XYS1 Length: 438 May 30, 2002 09:32 Type: P Check: 646 ..

1 MTSSDSKLAG MISSESIEIC EVKMLIEHS LMSRFHAFVN EMITVKNRR
 51 MFPVLTSTIT GLDPTAMYSV MLDFFVPVDNN RMKYVNGEMI PGKPEPHVS
 101 SCAYIHDPSP NFGSHMKOP IGFSRVKLTN KATGNPQOIM LNSLHKRPR
 151 IHKRVGAVE SQQVVAHSF QETRIAYTA YONEDVTSLK IKYNPEAKAF
 201 LDGKEERSGN ENYFKDSTKA GSSQVYPRAT TWFPNOSNPT FNOCQYEPGI
 251 PPFIYPIPNQ PRTTRRRMS RTQRSHPRKP PTQTGYODFQ PTNYPIIPTD
 301 QWSSISGHE LDEGHFSLEP VSVDDVTAIG FDTPHQAFAP NDLISIEPSY
 351 SLDPQFTTT WSSRHMPGY VNSPIRSLE RPEFYRGVD VSOQNYVTMT
 401 STDASVTSS LYETPSPGVI QRSQEDIQSV YAVTPPSL

11AA-SEQUENCE 1.0
 ID Q9GUZ8 PRELIMINARY; PRT; 334 AA.
 AC Q9GUZ8;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE BRACHYURY PROTEIN.
 GN ODT.
 OS Oikopleura dioica.
 OS Eukaryota; Metazoa; Chordata; Urochordata; Appendicularia;
 OC Oikopleuridae; Oikopleura.
 OX NCBI_TaxID=34765;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bassham S., Postlethwait J.H.;
 RT "Brachyury (T) expression in embryos of a larvacean urochordate,
 RT Oikopleura dioica, and the ancestral role of brachyury.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF204208; AAG22592.1; -;
 DR HSSP: P24781; 1XBR.
 DR InterPro: IPR001699; T-box.
 DR Pfam: PF00907; T-box; 1.
 DR PRINTS: PR00937; TBOX.
 DR SMART: SM00425; TBOX; 1.
 DR PROSITE: PS01283; TBOX_1; 1.
 DR PROSITE: PS50252; TBOX_3; 1.
 SQ SEQUENCE 334 AA; 37574 MW; 02AEA233B5800F50 CRC64;

Q9GUZ8 Length: 334 May 30, 2002 09:32 Type: P Check: 9657 ..

1 MSPLDIKIED APSEADQIEV KLEEPPLMK FSNLTNEIV TKNGRRFPV
 51 MKVETKGLDE NAMTYVLDF TAADNHRMKY VNGEVPCKP PEOVPESSV
 101 MHPDSPNGS HMKKPLSFS KYKLTKLNQ HNGQOIMLS LHKYEPVHI
 151 IKIGTAGAQ EFVKTPFPW TRFIAYTAQ NEETSLSKIR HNPFAKFLD
 201 AEORKQNDYL PISASANSN SNGRSAONN RKRRAENRTQ PYKRPVPAK

```

251 SESISSTDFG AQNMMPSTQ SMGIPIDAPA RTFSPDQSSA VLGDPLPAH
301 ESGSPPLSMP PMNDWCQNGY YSTGLEDAFL ELLV

11AA_SEQUENCE 1.0
ID Q9NHK6 PRELIMINARY; PRT; 110 AA.
AC Q9NHK6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
DE PFEM1 PROTEIN (FRAGMENT).
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP STRAIN=19;
RC SEQUENCE FROM N.A.
RX MEDLINE=20520970; PubMed=11071291;
RA Taylor H.M., Kyes S.A., Newbold C.I.;
RT "Var gene diversity in Plasmodium falciparum is generated by frequent recombination events.";
RL Mol. Biochem. Parasitol. 110:391-397(2000).
DR EMBL; AF221784; AAF36624.1; -.
FT NON_TER 1 110
FT SEQUENCE 110 AA; 12828 MW; 65ED0504CF943FA1 CRC64;

Q9NHK6 Length: 110 May 30, 2002 09:32 Type: P Check: 7692 ..

1 DLGDIIIRGKD LVGHKLGNN KLEARLQTMF QNIKNKNPP LDKLSLEKRR
51 EYWMALNRKE YMKAITCRAG ESDRYSKTIT YGTATSNKG CGHDDQDVOT
101 YLDYVPOFLR

11AA_SEQUENCE 1.0
ID Q9NHG3 PRELIMINARY; PRT; 110 AA.
AC Q9NHG3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
DE PFEM1 PROTEIN (FRAGMENT).
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RX MEDLINE=20520970; PubMed=11071291;
RA Taylor H.M., Kyes S.A., Newbold C.I.;
RT "Var gene diversity in Plasmodium falciparum is generated by frequent recombination events.";
RL Mol. Biochem. Parasitol. 110:391-397(2000).
DR EMBL; AF221827; AAF36667.1; -.
FT NON_TER 1 110
FT SEQUENCE 110 AA; 13061 MW; 1A30E3C41C607D35 CRC64;

Q9NHG3 Length: 110 May 30, 2002 09:32 Type: P Check: 7219 ..

1 DIDDIYGRD LFLGHKQKN ELEARLQMF ENIKENNGEL KKLLEQFRE
51 YWMALNRQV WKAITCAAD NDKYSKITDN GRITFSYDKC GHHVNDVPT
101 NLDYVPOYLR

11AA_SEQUENCE 1.0
ID Q9U218 PRELIMINARY; PRT; 569 AA.
AC Q9U218;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, last annotation update)
DE Y41C4A.2 PROTEIN.
GN Y41C4A.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae; Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Steward C.A.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AL032627; CAB6353.1; -.
DR InterPro; IPR000120; Amidase.
DR Pfam; PF01425; Amidase; 2.
SQ SEQUENCE 569 AA; 62918 MW; D0827E77B2A7958D CRC64;

Q9U218 Length: 569 May 30, 2002 09:32 Type: P Check: 2401 ..

1 MRRYKGGMHE ADRFOAGARP CLPTMEVLVS EINSQNNFEK VLIHQRTAPE
51 LQSDDETFE RRLISEQLL GDSRYLGKA TELDLDFLR AVYDDLQSL
101 IDKLQAKDGL NAYTVLCATA RRMDDCQTRL DCVAGVMEV GOVTSGARAF
151 QTAQDTDTLM YNSEKPPY GIPESIKCDG ESGIANSPT THLRMGAIIP
201 FVTTISIKTP HPRNPWALDC SPGLCGSEA ALVADGAPF GRAPDLDAAS
251 LRIASAFQGL VTLKPTRDRF HVSGSNSYGF YTKNVQDQVF LKLPFGSAG
301 YRGLEPNSSP APLMDKLEN KIKIGWFEDD GFNAPVPSNR RAVVDTIGLL
351 EKOGEYEVKF EMEIDFPFPQ VAQMEFFQQR PYDSNFIPNL YKSNQCLGR
401 FGKFLNPPL VARISKRAV ISQSNMKEM CKNIEDIESY KIKFIEYKWS
451 LGVDVLICPA FCIPAVPEKY LPELVNTRLS TGLENNLDFP AGIVPAHYT
501 ADVVANLEDE KIFPIDALL RKQRDACVNS EAMPNSVOIV GLPNEETVL
551 EVMKIVENLH GPMSNPKGF

11AA_SEQUENCE 1.0
ID Q9U2X4 PRELIMINARY; PRT; 331 AA.
AC Q9U2X4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, last annotation update)
DE Y11367B.9 PROTEIN.
GN Y11367B.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae; Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Leonard N.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for investigating biology.";
RL Science 282:2012-2018(1998).

```


DR EMBL; AL110477; CAB54331.1; -.
 DR InterPro: IPR002892; DUF40.
 DR Pfam; PF01838; DUF40; 1.
 SQ SEQUENCE 331 AA; 38135 MW; C2F670ACBEDA5462 CRC64;

Q9U24 Length: 331 May 30, 2002 09:32 Type: P Check: 4393 ..

1 MEIILQLL LGIISAITTI LNLINLVKI VLNPSKRND MYLFFYRFTL
 51 DIFFGAGLS YIAYTLNME APEFMFOYS LVLALPMS HISCRSIIA
 101 LSIISDRSIA TCFPIYFFKN RKKIPMPVL LIGSLGLAE EYMLFGCSY
 151 NMEIPKTCV FGCATNOCF HWLIORSII FSLIVFSLI LSIKLMMNS
 201 VKHOOSNOI SKANRLALD TCVLFEDFL PAFCGHMPT APMSFNWV
 251 SYNPLKLTG CAIBSTVTR VLLFRTPENS ASTPVOKLIY TLRPLSLAQ
 301 FEPPIKNIPE HFLRSRAVTV LYKNIKLIO F

11AA_SEQUENCE 1.0
 ID Q9N8U9 PRELIMINARY; PRT; 301 AA.
 AC Q9N8U9;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE POSSIBLE H-SCOL.
 GN CHRL138.
 OS Trypanosoma brucei.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TREU927;
 RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
 RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
 RA Gerard C., Rajandream M.A., Barrell B.G.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL359782; CAB95434.1; -.
 DR InterPro: IPR003782; SCOL_Senc.
 DR Pfam; PF02630; SCOL_Senc; 1.
 SQ SEQUENCE 301 AA; 33578 MW; F03C9365F6CBF559 CRC64;

Q9N8U9 Length: 301 May 30, 2002 09:32 Type: P Check: 7148 ..

1 MSLRHVITR SISVHNFIS SSGEQKTGGS RGFESVLRV GPRRAELFG
 51 LCLGACMPL SIYLNMCCEG CSNRNNNTOP ELVLGNETFV THGVAAPDKG
 101 SKYPLGCPFR LRESRTGNYI TDKELFQDHW TLLYFGFSKC AEVCPSTLRF
 151 ITDVAKADE KLAGDKMIST EAARLAVFL SVDSRRDTE VLEGVSKYD
 201 PRVRLGTGS KEIQOAAAW RYVYSSIDET DEKSAAREAK GVPWAGADD
 251 TYOLDHSSAI YLVGVGDKLK DFEFKMGVA DAVGRLEVL QDVGFKDTR
 301 G

11AA_SEQUENCE 1.0
 ID Q9N8E6 PRELIMINARY; PRT; 136 AA.
 AC Q9N8E6;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE HYPOTHEICAL 16.2 KDA PROTEIN.
 GN CHRL136.
 OS Trypanosoma brucei.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5691;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-TREU927;
 RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
 RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
 RA Gerard C., Rajandream M.A., Barrell B.G.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL359782; CAB95597.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 136 AA; 16200 MW; CACDC1CE14964C54 CRC64;

Q9N8E6 Length: 136 May 30, 2002 09:32 Type: P Check: 5512 ..

1 MCQSTKSN VEAPASEMKN KERRKEKER ECGKYLKKEK NTKRESFSL
 51 FFKMLFTLL FHEFWFICL FVSFSEFLLS PDRNSVPLTV YIMATVPQY
 101 VTIATTVIIL YIYIYIYIIL KITTFITTSF YIYIHC

11AA_SEQUENCE 1.0
 ID Q94216 PRELIMINARY; PRT; 589 AA.
 AC Q94216;
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE HYPOTHEICAL 69.4 KDA PROTEIN.
 GN F38A5.13.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RN [1]
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium. ";
 RL Science 282:2012-2018 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA "Fulton B., Stellyes L.;
 RT "The sequence of C. elegans cosmid F38A5. ";
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA "waterston R.;
 RT "Direct Submission. ";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U70854; AAB09157.1; -.
 DR InterPro: IPR001623; DnaJ_N.
 DR InterPro: IPR001005; Myb_DNA_bind.
 DR Pfam; PF00226; DnaJ; 1.
 DR Pfam; PF00249; Myb_DNA-binding; 2.
 DR SMART; SM00271; DnaJ; 1.
 DR SMART; SM00395; SANT; 2.
 DR PROSITE; PS00976; DNaJ_2; 1.
 DR PROSITE; PS50090; MYB_3; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 589 AA; 69413 MW; D3CF39ED6B3EED9 CRC64;

Q94216 Length: 589 May 30, 2002 09:32 Type: P Check: 5664 ..

1 MTGMLQVAI YGTFARSAF EPAGICETR LIRDKITLGH CTLPYIKNSP
 51 VAKPYTEQOR KKSSEPAEL RKDLFDADNE KYEKYLMKID PNDCKNDHY
 101 KYLGSLKLRW QATSDAIRFC YRQVKLKHNP DKKHNGIYW EKEEYFCIT
 151 KAYEYQMSD VKRAFDSDVD HKENDIIPNE KSIINNNFN ELAPYFOINS
 201 RMSNIKPYPE LKGSADATRED VENFYDFWPN FOSWREFSYL DEEDKERGED

```

251 RYERREMEKO NKAERERRR EAARIRKIV DIAAKDPRI IKRKKEQAK
301 KDKAKEDKOR AIREKQEAID REKREKEAE AKOKEADRK AKERERERKK
351 ERDIARRAMS QQRRLKLIA DEAGHWTENP RDKLTMEMRI ERICIGFTVD
401 QLRLECKEVE SLSIASEIOT ALTDAILLKK EAAGAKYIT EDKNKEKEKO
451 ADKETWTSEE IQLLVASNT FPPGTERVW QIADYINEHR KDSTGLPPKT
501 EKOYIKOCKA VOTMVKLPS TTQNLGTAL PDEDWMSATE OKTLEDAIKK
551 HKSSDPERWE KISTEVGTGS KKACIRRRKY LVQVAKNKK

11AA_SEQUENCE 1.0
ID 046218 PRELIMINARY; PRT; 219 AA.
AC 046218;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE GLUTATHIONE S TRANSFERASE-1.
GN CUGSTL.
OS Culicoides varilipennis.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Prexygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Ceratopogonidae; Culicoides.
OX NCBI_TaxID=46212;
RN [1]
RP SEQUENCE FROM N.A.
RA Abdallah M.A., Pollenz R.S., Tabachnick W.A., Nunamaker R.A.,
RA Murphy K.E.;
RT "Identification of a cDNA clone encoding Culicoides varilipennis
RT glutathione S-transferase (GST).";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U087958; AAB94639.1; -.
DR HSSP; P30712; ILJR.
DR InterPro; IPR004046; GST_C.
DR InterPro; IPR004045; GST_N.
DR Pfam; PF00043; GST_C; 1.
DR Pfam; PF02798; GST_N; 1.
KW Transferase.
SQ SEQUENCE 219 AA; 24769 MW; A9EBD80ACCB157 CRC64;
046218 Length: 219 May 30, 2002 09:32 Type: P Check: 7893 ..

1 MGDEFTLPG SSEPCRAVOMT AKAVGVDLNL KLTNLMAGEH MKDEFKLNP
51 QHCIPTLVDN GFSLWESRAI QVLYEKYCK DDSLPRKDVQ QRALVNOQLY
101 FDMGTLYORF ADVWYPOLFA KQAPNPNFK XMEEMGFNL TTEGCHKYAV
151 GDKETVADLIA LAASVATVEY SGDFKPPYN VQKWFALCKT TLDGYDLNEA
201 GVKNSRIPL SLNACSLN

11AA_SEQUENCE 1.0
ID 017999 PRELIMINARY; PRT; 814 AA.
AC 017999;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE C15A7.1 PROTEIN.
GN C15A7.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoloidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z67735; CAA91530.1; -.
DR InterPro; IPR001175; Neur_chan.
DR Pfam; PF02931; Neur_chan_LBD; 2.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; UNKNOWN_2.
SQ SEQUENCE 814 AA; 95185 MW; 36464FAB9017B51D CRC64;
017999 Length: 814 May 30, 2002 09:32 Type: P Check: 7776 ..

1 MIFPLMILVL FNTPIRVSE ADESEERHOA YLDHRKLKED VLQPLDGSMP
51 SELPVGIGAQ LKLDHISRDV EKTGFLIAL TVKRWMSDER LQMDPSHYSG
101 FTRRLREVD FWLYQWKRK IYMTTYERK QSKTIDLNL TLVEMDIHKK
151 GFWVTTTKIL LKTECFIDRK GYPHDFONCS FSLIPNNAN DIFLGSSMLK
201 AKFKRDEFTDP QQVVRVQDFQ INDVTEDLX LFNRIYLHD LTRPRSLVR
251 SFVSFTVLQ RKNVYVHLD YAPMFICTC LIISGLPSE YALPLVFNL
301 ISEYLYPHNI KDVLPTDFDG TPTIALIATF TLAFTMSIIG WKMFITTIMT
351 KKQKAYFLNS SKRPLDSGFF NIRVLRLILT IDKFLMYMLI AQSGFKPTTI
401 FLQLEFFQRA ELRMLLKYNA RQNPFPVLAS QYSKINISF SLEESSYNS
451 YFHCARYPS SCGVTMGNGK KLHGFDHGG FMSVYINAEI VSYLYVQVE
501 SFIFQWIDE RLKMNPNANS GYREIVEKTF EFRKDNCCMW PIKYRSYDR
551 RYSELDLLEF SDARTLISYK GEIKALQTM VTTQCQSFQ EYENDQONS
601 IMLIPNOMAD EFRFVSPNGY CNPKFENLEH RAVRVHDIHL MGVESNIEYT
651 FANTYFTAAE VTGEFYMAKT EFRFNLMEKR VNKLINXLS IPSILSMFL
701 IAGLFPNGY SIFGVAYCFE VEVHMGLYVS KIIPNDIRGI PYGALALCE
751 LIETVLEFCW KVPSTYARNQ KLEQLPNVF ADNRFVYIKP VEVYDRVCV
801 ALCLQFLSVI YKTR

11AA_SEQUENCE 1.0
ID 018843 PRELIMINARY; PRT; 660 AA.
AC 018843;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE C54G10.3 PROTEIN.
GN C54G10.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoloidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthews L.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z75532; CAA99810.1; -.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001687; ATP_GTP_A.

```

DR Pfam; PF00005; ABC_tran; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
SQ SEQUENCE 660 AA; 74999 MM; 0A59EB7C8C5079CA CRC64;

Q18843 Length: 660 May 30, 2002 09:32 Type: P Check: 7424 ..

1 MARLALDFLF FRRFRWLKLI LPPCHRRSIT AMLAGALLIA TALDQVMTYL
51 VGIPSLFEYV ALGNKADATF KILCARGAAI ILGKAFITLAV FKYLITNLAI
101 KSRQVCNLTML HRLYFKRQAF FKLSGSDML DNPDRQLTOD IEKATRLISN
151 DLAPLAPAF FIIGYTWLT YESSGWIGPA AIYFFCIOT IINKMILSPI
201 VQKVEDEKM EGDPRORHME VRSNVEALAF YRAGLENIM TNOKLNLIE
251 TQKSLTEKRM VLNSITNVED YEGGILSYLI IGVPIFTHL YDDVSPLELN
301 GIVSRNAFFY LYLIFSSTV LKLTGDFGEL AGVTRHVMEL HEELNRLHSD
351 CLETRDRPST VPSSVYVIAS DEDDKSASRH MQEIHGKQMS LERDEQEEEE
401 AQLYLGCTG QEDDWPDDGV AITVDSATLS PPNDHSHLIY QLSLQIITQG
451 QTLITGDSG CGKSSILRMP AGLMHCSGK MDCHMRRLTS NLFLAOKPY
501 FPGSGTTLRQ QIYYPVKALQ VDKDVARITQ ILEWVKMEHL VERGGGLDTP
551 VEMDMKLTLS PGELQRLSLA RVFYTKPRIV FLDESTSAIG FELEMAIYRK
601 LQEKITFVS IGHRYSLKQF HDMELRVKGR SGEWSLHDID TASIASRTAS
651 FLGADTVLSM

11AA_SEQUENCE 1.0
ID Q9XUC9 PRELIMINARY; PRT; 170 AA.
AC Q9XUC9;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE M162.2 PROTEIN.
GN M162.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Basham V.M.;
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
investigating biology."; Science 282:2012-2018(1998).
RL EMBL: 282278; CAB05254.1; -.
DR EMBL: 282278; CAB05254.1; -.
DR InterPro: IPR001304; Lectin_c.
DR Pfam; PF00059; Lectin_c; 1.
DR SMART; SM00034; CLCCT; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 170 AA; 19234 MM; 1EB4956690519F17 CRC64;

Q9XUC9 Length: 170 May 30, 2002 09:32 Type: P Check: 6239 ..

1 MFPKLLFLAI LIPISSLTDCP DIDGQTRERK GICYKFTVSL MKYDARDWC
51 HHNNPAGPSW IAYVPOETS NFLAYVAGSI FGEQFKHFWI GLNDRPISKT
101 LSWDTGLSVS YTNFGSNVAQ NTFSENITNT KWNITGDDDEV HNFVCSRPS
151 TVPATVTRQAP QARRLAAMKN

11AA_SEQUENCE 1.0
ID O45282 PRELIMINARY; PRT; 705 AA.
AC O45282;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE C3IA11.5 PROTEIN.
GN C3IA11.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurtry A.A.;
RL Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
investigating biology."; Science 282:2012-2018(1998).
RL EMBL: 283218; CAB05687.1; -.
DR InterPro: IPR003862; UPF0063.
DR Pfam; PF02712; UPF0063; 1.
SQ SEQUENCE 705 AA; 81206 MM; 03C65ACADD52F9D8 CRC64;

O45282 Length: 705 May 30, 2002 09:32 Type: P Check: 676 ..

1 MKIAPLFEFI VCLPATEFCE NMKEVFIARA ADOKITGVSE QCANDTETWQ
51 KSKMVAEELS AECLIEOKCT KEELKTIEDN FVAVEQYDAM GKILTGELFO
101 LPILEDGSYQ ECEKISGSKY ATNYCYWVLM PCKNATCHMS DGLPTTFEER
151 GAVCPYSCS EODLPVYVQD VSDPPTACA APCSSYPVKK TPAPMGFTSF
201 MAMVMGIML ATYDIYAKDA LKREDEKRED SKILDIILTF SLMTNAELL
251 SVKEOKPGFI KCLDCIRFLS MLMVVTGHTF SYLTTPDOIE SILPFGFRFW
301 NMLVNAFYS VDPFLLSGL VVSYLEFKTK LKVSQIKSPI TWILFYVHR
351 LRLTPPLMEF LGFEVYVYGY FQGPVYASOL NOQNEVDTC QTYMKNLIIY
401 INNLMSGDTQ CYGITWYIGA DTQLYLVAPI FLIGLFPSEA IGTALLTAAAT
451 IGSVITVYIL FSTYDLPADF FGNGDATTHFY DMIIYIKPMIR CPYFVGILV
501 GYLATYGRK KLRUNMALAV TGMIVAFSLG ALCITSTIDY DNKKWSIFTS
551 RAYVYNSRL AMSFALSMTI VANHMGWGP IDAFNSHPMW QPFGRLSYCA
601 YIVHYVVLVM YLMIGDASIH FYSSPQIFW YAVPTVLSY IFAFWMSOLF
651 EEPFLKLEKM LIELITGAR DRNREDIEKO KILLTKENE LMAVEEYVOS
701 TNEKF

11AA_SEQUENCE 1.0
ID O01616 PRELIMINARY; PRT; 533 AA.
AC O01616;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE HYPOTHETICAL 60.8 KDA PROTEIN T19H12.10 IN CHROMOSOME V.
GN T19H12.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Davidson S.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO FAMILY UPF.
DR EMBL: U97009; AAC69033.1; -.
DR WormPep: T19H12.10; CE13770.
DR InterPro: IPR002213; UDPGT.
DR Hypothetical protein; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT TRANSMEM 499 519 POTENTIAL.
FT SEQUENCE 533 AA; 60763 MW; E4F111CFC067124 CRC64;

001616 Length: 533 May 30, 2002 09:32 Type: P Check: 9544 ..

1 MRLTFSCLC FLKVNVAAK ILIENPIGF SHKVFVSKLA DIIADHGQEV
51 TLFQPFHLAL KNLGGIVKKN NIEINLYPD HYDDLKLET QTFPDEWDSQ
101 LMNPVLMAF MLPRILGGEF KKTIIQLKD KEILKLIKDK KPDVAISETF
151 ELIGMYSHF LGVPCIPILS AVRLDIFNEA FGSSAFGYL TQGSKLAPD
201 AGFLDLRLNDV YRDFESKMAF RGMAYQNDV IEKAAGHVP YMKDLVKEAP
251 VYMTNSPYL DPAVPTTATI VHIGITINL EKMNVADLP EEVEIILKEK
301 ETVVLISFGS VIRSYEMPEN EKAGLIKVE SLDPVITIKW YEIDLEFPK
351 KLRPNVHLAK WVPQPSLAD KRVKLEVTNG GLGSTMENVAY TGRPALMVP
401 FGDQPMNAD LARHGGAIAV DKFDLVGKK LLETVRDLVT NPKYEQAKE
451 LDVLTNQPI DRYMNMKHL EFAIKFPMLR SQPEINQV PIHYYLDVI
501 VFLFVSIT ATISFOIVCR ILSRLSKV KSD

11AA_SEQUENCE 1.0
ID 004622 PRELIMINARY; PRT; 326 AA.
AC 004622;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE U2 SMALL NUCLEAR RIBONUCLEOPROTEIN 40K.
OS Trypanosoma brucei.
OC Eukaryota; Euzoenzoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5702;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=427;
RX MEDLINE=93209224; Pubmed=8458336;
RA Cross M., Wieland B., Ralfi Z., Guenzl A., Roethlisberger U.,
RA Lahm H.-W., Binderl A.;
RL The trans-splicedomal U2 snRP protein 40K of Trypanosoma brucei:
RT cloning and analysis of functional domains reveals homology to a
RT mammalian snRP protein";
RL EMBL: J. 12.1239-1248(1993).
CC -1- FUNCTION: THIS PROTEIN IS ASSOCIATED WITH SN-RNP U2. IT BINDS STEM
CC LOOP IV IN COMBINATION WITH OTHER TRYPANOSOMAL PROTEIN(S).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
CC MANY PROTEINS. NUMBER IN THIS PROTEIN: 7
CC -1- SIMILARITY: TO THE N-TERMINAL HALF OF THE HUMAN U2 SNRNP-A'
CC PROTEIN.
DR EMBL: X69934; CAA49552.1; -.
DR HSSP: P09661; I49N.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat;
KW Leucine-repeat.

```

```

FT DOMAIN 21 215 LEUCINE-RICH REPEATS.
FT REPEAT 21 33 LRR 1.
FT REPEAT 34 58 LRR 2.
FT REPEAT 84 101 LRR 3.
FT REPEAT 102 126 LRR 4.
FT REPEAT 127 151 LRR 5.
FT REPEAT 152 176 LRR 6.
FT REPEAT 190 215 LRR 7.
FT MUTAGEN 1 20 MISSING: IN 40K-1; REDUCES U2 BINDING.
FT MUTAGEN 154 326 MISSING: IN 40K-2; ABOLISHES U2 AND
FT MUTAGEN 189 326 GENERAL RNA BINDING.
FT MUTAGEN 215 326 MISSING: IN 40K-3; ABOLISHES U2 AND
FT MUTAGEN 215 326 GENERAL RNA BINDING.
FT SEQUENCE 326 AA; 36595 MW; 0B42A08A705C1B6A CRC64;

004622 Length: 326 May 30, 2002 09:32 Type: P Check: 3173 ..

1 MRLTDTIRR ADQFTNALRO RELDRLGLGI TVEEHTLTF LNDSPVLNL
51 SQNPRLLEY FPGDSAPLAT AAQNSSAKP ASRWMLRLQT LVYHRRRLTH
101 VSEATCAVL PNIRAFVADH NEFRLRLD LLSHWKLEI LSTENHPITI
151 SEDNARILRAY VVFLCPTLKL VNYQRYTOVD RQNYETMRKE FVGLVEGWRR
201 LEAKQLLOQN SAPTEADASA SESYKIRKR SRHAREASK NGSADTAPEP
251 AASVETFEAT PSYAAEAGE EASNALQAR LEALEEKMAA AETEEELMEL
301 QQELTELET MKHQASKGT KKTPTS

11AA_SEQUENCE 1.0
ID 095JP5 PRELIMINARY; PRT; 235 AA.
AC 095JP5;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MACACA fascicularis (Crab eating macaque) (Cynomolgus monkey).
OS Macaca fascicularis (Primates; Catarrhini; Cercopithecoidea;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RA Terao K., Sugano S.;
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
RT libraries.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB070136; BAB63081.1; -.
KW Hypothetical protein.
SO SEQUENCE 235 AA; 24993 MW; D5D728E1149BD3F0 CRC64;

095JP5 Length: 235 May 30, 2002 09:32 Type: P Check: 5683 ..

1 MNSTATTH SCPIILVPS YWASRSCMW TRLVTPICIL LPTQCEDE
51 DENLYDPL LNNPEARSS NSKTTTTLGL VVHAADGVA LGAAASTQST
101 SVOLIVFAI MLHKAPAFG LVSFIMHGL ERNRIRKHL VESLAPVMS
151 MYVYLGLSKS SKEALSEVNA TGMALFESAG TFLVAVVHV LPEVIGIGHS
201 HKPDATGGRG LSRLEYVALV LGCLIPILIS VGHQH

11AA_SEQUENCE 1.0
ID 09TGJ4 PRELIMINARY; PRT; 161 AA.
AC 09TGJ4;
DT 01-MAY-2000 (Tremblrel. 13, Created)

```

DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE NADH DEHYDROGENASE SUBUNIT 4 (FRAGMENT).
 GN NDA.
 OS Taenia crassiceps.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
 CC Cyclophyllidae; Taeniidae; Taenia.
 OX NCBI_TaxID=6207;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21036605; PubMed=11163447;
 RX Nakao M., Sako Y., Yokoyama N., Fukunaga M., Ito A.;
 RT "Mitochondrial genetic code in cestodes."
 RL Mol. Biochem. Parasitol. 111:415-424(2000).
 DR EMBL; AB031287; BAA83540.3; -.
 KW Mitochondrion.
 FT NON_TER 1
 SQ SEQUENCE 161 AA; 18553 MW; 4FE537849D6CBDE8 CRC64;
 O9TCJ4 Length: 161 May 30, 2002 09:32 Type: P Check: 134 ..

1 AFLSHSHVY PFLCFTSDM SLSLSEFPG MGHGASAGV FCLLMCFYDI
 51 SNTRMWLIK SSSINGIMIM IVFSLSLIC SPTTIQFPC EYMLVINSN
 101 LLLLIIFWC YLFLGLVPL VLGCHLLRS EYESVGVSC YSYFFLVFS
 151 CPWCLGFFI F

11AA_SEQUENCE 1.0
 ID Q9B8W7 PRELIMINARY; PRT; 417 AA.
 AC Q9B8W7;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 4 (EC 1.6.5.3).
 GN NAD4.
 OS Taenia crassiceps.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
 CC Cyclophyllidae; Taeniidae; Taenia.
 OX NCBI_TaxID=6207;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AMERICAN;
 RX MEDLINE=20349913; PubMed=10869225;
 RA Le T.H., Blair D., Agatsuna T., Humair P.F., Campbell N.J.,
 RA Iwagami M., Littlewood D.T., Peacock B., Johnston D.A., Bartley J.,
 RA Rollinson D., Heriot E.A., Zarling D.S., McManus D.P.;
 RT "Phylogenies inferred from mitochondrial gene orders-a cautionary tale
 from the parasitic flatworms";
 RL Mol. Biol. Evol. 17:1123-1125(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AMERICAN;
 RA Blair D., Le T.H., Littlewood D.T.J., Zarling D.S., McManus D.P.;
 RL Submitted (DNC-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
 DR EMBL; AF216699; AAG13172.2; -.
 DR InterPro: IPR003918; NADH-oxdrctase4.
 DR Pfam: PF00361; oxidored_q1; 1.
 DR PRINTS: PR01437; NHOXDRDTASE4.
 KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
 SQ SEQUENCE 417 AA; 48714 MW; 521EF23852326BC CRC64;
 O9B8W7 Length: 417 May 30, 2002 09:32 Type: P Check: 7748 ..

1 MFIFFSVNF IVLVLFVLL LVFSCGVNCF SMINSVYNG LFVDSIFFY
 51 LFLVLFLGL YSNVMTNLL SPTRYLVLY SLFTSLSCF INHCVFENC

101 YELSMPLLY LIFCDHSHSE RFLAGWYFCS YLLITSPLI LILLYSVN
 151 NSFYSSMYA GDDVYLLVY LLSFVFETKV PLVPHTMLP IVHAEATSIY
 201 SIFLSGYIMK LGLGVRCS CFIFNVSFLM YLFICICILCV FFLVSCSEL
 251 DGRMLAFLS LSHIVPELC LFISDSSLS FSPFYCMGHG LSAIVFCCL
 301 WCFYDISNTR NWVLKRSIN GIMLMIVFE SLSLCSPT TIQFCEVNL
 351 VINSSNLLL IIFWFCYFL GGLVPLVLCG HLLINSEFE SVGVCSTSYF
 401 YELVSCFCWC YLGFIF

11AA_SEQUENCE 1.0
 ID Q94TE9 PRELIMINARY; PRT; 116 AA.
 AC Q94TE9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE NADH DEHYDROGENASE SUBUNIT 3.
 GN ND3.
 OS Chlorophthalmus agassizii (shortnose greeneye).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OX Autloptiformes; Chlorophthalmidae; Chlorophthalmidae; Chlorophthalmus.
 OX NCBI_TaxID=143313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miya M.;
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21519020; PubMed=11606696;
 RA Miya M., Kawaguchi A., Nishida M.;
 RT "Mitogenomic exploration of higher teleostean phylogenies: A case
 study for moderate-scale evolutionary genomics with 38 newly
 determined complete mitochondrial DNA sequences."
 RL Mol. Biol. Evol. 18:1993-2009(2001).
 DR EMBL; AF002918; BAB70021.1; -.
 KW Mitochondrion.
 SQ SEQUENCE 116 AA; 12978 MW; 10657EE65DFA5AB CRC64;
 O94TE9 Length: 116 May 30, 2002 09:32 Type: P Check: 2807 ..

1 MNLVSTVIMT ALAISLLAF VSWLPOLNP DTEKLSPTFC GFDPGSAHL
 51 PSLRFFLYA ILFLFLDEI ALLPLPWGD QLTPTHTFY WATAVLALLT
 101 LGLAVEMTQG GLEWAE

11AA_SEQUENCE 1.0
 ID Q94TE0 PRELIMINARY; PRT; 116 AA.
 AC Q94TE0;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE NADH DEHYDROGENASE SUBUNIT 3.
 GN ND3.
 OS Trachipterus trachipterus (ribbon fish).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OX Acarctomorphia; Lampridiformes; Trachipteridae; Trachipterus.
 OX NCBI_TaxID=143894;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miya M.;
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21519020; PubMed=11606696;
 O94TE0 Length: 116 May 30, 2002 09:32 Type: P Check: 2807 ..

RA Miya M., Kawaguchi A., Nishida M.;
 RT "Mitogenomic exploration of higher teleostean phylogenies: A case
 study for moderate-scale evolutionary genomics with 38 newly
 determined complete mitochondrial DNA sequences.";
 RL Mol. Biol. Evol. 18:1993-2009(2001).
 DR EMBL: AP002925; BAB70112.1; -
 KM Mitochondrion
 SQ SEQUENCE 116 AA; 13135 MW; 020ED948A2E92E93 CRC64;
 Q94N60 Length: 116 May 30, 2002 09:32 Type: P Check: 3118 ..
 1 MNLVTSLLI ALTLSTVLIF ISFWIPQLSP DHEKLSPEEC GFDPILGSARL
 51 PFSLRFFLVA ILFLFDLEI ALLPLPMGN HLASPLEFTT WATLVLLILT
 101 LGLIYEWYQG GLEWAE
 11AA SEQUENCE 1.0
 ID Q94T47; PRELIMINARY; PRT; 116 AA.
 AC Q94T47;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE NADH DEHYDROGENASE SUBUNIT 3.
 GN ND3.
 OS Zu cristatus (scallop ribbonfish).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Lampridiformes; Trachipteridae; Zu.
 OX NCBI_TaxID=143325;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miya M.;
 RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RX MEDLINE-21519020; Pubmed-11606696;
 RA Miya M., Kawaguchi A., Nishida M.;
 RT "Mitogenomic exploration of higher teleostean phylogenies: A case
 study for moderate-scale evolutionary genomics with 38 newly
 determined complete mitochondrial DNA sequences.";
 RL Mol. Biol. Evol. 18:1993-2009(2001).
 DR EMBL: AP002926; BAB70125.1; -
 KM Mitochondrion.
 SQ SEQUENCE 116 AA; 13145 MW; E8E24CB7AAE295F9 CRC64;
 Q94T47 Length: 116 May 30, 2002 09:32 Type: P Check: 2171 ..
 1 MNLVTSLLI ALTLSTVLIF ISFWIPQLSP DHEKLSPEEC GFDPILGSARL
 51 PFSLRFFLVA ILFLFDLEI ALLPLPMGN HLACPDFTFI WATLVLLILT
 101 LGLIYEWYQG GLEWAE
 11AA SEQUENCE 1.0
 ID Q94S17; PRELIMINARY; PRT; 116 AA.
 AC Q94S17;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE NADH DEHYDROGENASE SUBUNIT 3.
 GN ND3.
 OS Antigonla capros (deepbody boarfish).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Zelformes; Caproidae; Antigonla.
 OX NCBI_TaxID=143339;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miya M.;
 RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21519020; Pubmed-11606696;
 RA Miya M., Kawaguchi A., Nishida M.;
 RT "Mitogenomic exploration of higher teleostean phylogenies: A case
 study for moderate-scale evolutionary genomics with 38 newly
 determined complete mitochondrial DNA sequences.";
 RL Mol. Biol. Evol. 18:1993-2009(2001).
 DR EMBL: AP002943; BAB70346.1; -
 KM Mitochondrion.
 SQ SEQUENCE 116 AA; 13055 MW; 2A471321E3C3291B CRC64;
 Q94S17 Length: 116 May 30, 2002 09:32 Type: P Check: 2075 ..
 1 MNLVTTIAMI SITLSTVLAF VSFWLPQWMP DHEKLSPEEC GFDPILGSARL
 51 PFSLRFFLVA ILFLFDLEI ALLPLPMGD QLSSPLLTLFL WAAAVLTLLLT
 101 LGLIYEWYQG GLEWAE
 11AA SEQUENCE 1.0
 ID Q94SD5; PRELIMINARY; PRT; 116 AA.
 AC Q94SD5;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE NADH DEHYDROGENASE SUBUNIT 3.
 GN ND3.
 OS Dactyloptena peterseni (starry flying gurnard).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Scopaeiformes;
 OC Dactylopteroidae; Dactylopteridae; Dactyloptena.
 OX NCBI_TaxID=143342;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miya M.;
 RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RX MEDLINE-21519020; Pubmed-11606696;
 RA Miya M., Kawaguchi A., Nishida M.;
 RT "Mitogenomic exploration of higher teleostean phylogenies: A case
 study for moderate-scale evolutionary genomics with 38 newly
 determined complete mitochondrial DNA sequences.";
 RL Mol. Biol. Evol. 18:1993-2009(2001).
 DR EMBL: AP002947; BAB70398.1; -
 KM Mitochondrion.
 SQ SEQUENCE 116 AA; 13092 MW; 40C800F75E078634 CRC64;
 Q94SD5 Length: 116 May 30, 2002 09:32 Type: P Check: 3914 ..
 1 MNLVTTIAMI SALLTIVLAF VSFWLPQWMP DHEKLSPEEC GFDPILGSARL
 51 PFSLRFFLVA ILFLFDLEI ALLPLPMGD QLSPVTTFL WASSVILLLT
 101 LGLIYEWYQG GLEWAE
 11AA SEQUENCE 1.0
 ID Q96303; PRELIMINARY; PRT; 534 AA.
 AC Q96303;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DE PHOSPHATE TRANSPORTER.
 GN PRT2 OR PRT4 OR PRT6.11.
 OS Arabidopsis thaliana (mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RX MEDLINE-97045083; PubMed-8927627;
 RA Muchhal U.S., Pardo J.M., Raghothama K.G.;
 RT "Phosphate transporters from the higher plant *Arabidopsis thaliana*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10519-10523(1996).
 RN [12]
 RN SEQUENCE FROM N.A.
 RC STRAIN-CV, LANDSBERG;
 RA Mbegule-A-Mbegule D., Gomez R.-M., Fils-Lycaon B.;
 RL Plant Physiol. 115:1288-1288(1997).
 RN [13]
 RN SEQUENCE FROM N.A.
 RC STRAIN-CV, LANDSBERG;
 RA Muchhal U.T., Muchhal U.S., Raghothama K.G.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [14]
 RN SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RX MEDLINE-99087485; PubMed-9872450;
 RA Okumura S., Mitsuoka N., Shirano Y., Shibata D.;
 RT "Phosphate transporter gene family of *Arabidopsis thaliana*.";
 RL DNA Res. 5:261-269(1998).
 RN [15]
 RN SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
 RT "Arabidopsis thaliana chromosome II BAC T7F6 genomic sequence."
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: U62331: AAB17266.1: -;
 DR EMBL: AF022872: AAB88291.1: -;
 DR EMBL: AB016166: BAA34398.1: -;
 DR EMBL: AC005770: AAC79607.1: -;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.
 KM SEQUENCE 534 AA; 58599 MW; 8186DBBCB8379E CRC64;
 SO

096303 Length: 534 May 30, 2002 09:32 Type: P Check: 605 ..

1 MAREOLOVLN ALDVAKTQWY HPTAIIAGM GFTDAYDLF CISLTKLIG
 51 RIYHVEGAQ KPGTLPPNVA AAVNGVAFGC TLAGOLFPGM LGDKLGRKKV
 101 YGWTLMVNL CSLSGLSFG HEPRKAVWATL CEFREWLGFSG IGGDYPLSAT
 151 IMSEYANKKT RGAFVSAVFA MGFEGIMAGG IFALIISAF EAKPSPAYVA
 201 DDALGSTITP ADLVWRITLM AGAIPAMTY YSRSKPETA RYTLAVADA
 251 KOASDSKRV LQVLEIPEEQ KLEISKEKS KAFGLEFSKEP MSRHGLHLG
 301 TTSTWFLDID AFYSQNLFOK DIFSAIGWIP PAQSMNAIOE VFKIAQOTL
 351 IALGSTVPGY WFTVAFTDVI GFAIQMGF FFMVTMFMAL ALPYHMHK
 401 ENRIGFVIMY SLTFEPANFG PNATTPVVA EIPPARFNST CHGISAASGK
 451 LGAMVGARFG LYLAQNPDKD KTDAGYPPGI GVRNSLIIVG VVNFGLILFT
 501 FLVPSKSGS LEEMSGENED NENSNDSTRT VPIV

11AA_SEQUENCE 1.0
 ID 050040 PRELIMINARY; PRT; 542 AA.
 AC 050040;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PUTATIVE PHOSPHATE TRANSPORTER.
 GN PHT5 OR T21L14.23.

OS *Arabidopsis thaliana* (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA Okumura S., Mitsuoka N., Shirano Y., Sato S., Shinozaki K.,
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA Rounsley S.D., Lin X., Kaul S., Ketchum K.A., Crosby M.L., Brandon R.C.,
 RA Sykes S.M., Kaul S., Mason T.M., Kellavagge A.R., Adams M.D.,
 RA Somerville C.R., Venter J.C.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: AB000093: BAA24280.1: -;
 DR EMBL: AC003033: AAB91982.1: -;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.
 KM SEQUENCE 542 AA; 59213 MW; 9B63C154899A8AC1 CRC64;
 SO

050040 Length: 542 May 30, 2002 09:32 Type: P Check: 992 ..

1 MARRKREVLN ALDAKTQWY HPTAIVAGM GFTDAYDLF SISLTKLIG
 51 RIYHVDSSK KPGTLPPNVA AAVNGVAFGC TLAGOLFPGM LGDKLGRKKV
 101 YGTTLMVNL CSLSGLSFG HSANGVWATL CEFREWLGFSG IGGDYPLSAT
 151 IMSEYANKKT RGAFIAVFA MGFEGILAGG IVSLIVSSTP DHAKAPYE
 201 VDPGSTVPGY ADYVWRIVLM FGAIPALITY YRMKMPETA RYTLAVARNT
 251 KOASDSKRV LQVLDIAEE AQSNSSNSP NTFGLFTPE FARRIGHLIL
 301 GTTTFWFLD IAYSSNLFQ KDITYALGTW PAETWMAIH EVFTYKAQT
 351 LIALGTVPG YWFTVAFTDI LGRFFIOLMG FIFMTFMPA LAIPDHRH
 401 RNRIGFELIM YSLMFRANF GPNATTPVP AELFPARLRS TCHGISAASG
 451 KGAIVGARF FLVAAQSSDS EKTDAGYPPG IGVNSLML ACVNFGLIVF
 501 TLVPSKSGK SLEISREDE EQSGGDTIVE MTVANGSRKV PV

11AA_SEQUENCE 1.0
 ID 096302 PRELIMINARY; PRT; 524 AA.
 AC 096302;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PHOSPHATE TRANSPORTER.
 GN PTL OR PHT1.
 OS *Arabidopsis thaliana* (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RX MEDLINE-97045083; PubMed-8927627;
 RA Muchhal U.S., Pardo J.M., Raghothama K.G.;
 RT "Phosphate transporters from the higher plant *Arabidopsis thaliana*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10519-10523(1996).
 RN [2]
 RN SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA: Pubmed-9192698;
 RA Mitsuoka N., Okumura S., Shirano Y., Sato S., Kato T., Harashima S.,
 RA Shibata D.;
 RT "Overexpression of an Arabidopsis thaliana high-affinity phosphate
 transporter gene in tobacco cultured cells enhances cell growth under
 phosphate-limited conditions";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:7098-7102(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA:
 RX MEDLINE-20181125; Pubmed-10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 features of the regions of 3,076,755 bp covered by sixty pl and TAC
 clones";
 RL DNA Res. 7:31-63(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: U62330; BAB17265.1; -;
 DR EMBL: D86608; BAA21504.1; -;
 DR EMBL: D86591; BAA21503.1; -;
 DR EMBL: AB025638; BAA97414.1; -;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 DR Transmembrane.
 KW SEQUENCE 524 AA; 57616 MW; 2D12D50987127BD6 CRC64;
 Q96302 Length: 524 May 30, 2002 09:32 Type: P Check: 8746 ..
 1 MAEQDLGVK ALDVAKTQL HFTAIYIAGM GFPTDAYDLF CVSLVTKILG
 51 RIYFNPESA KPSLPPHVA AAVNGVALCG TISGOLFPGW LGRKLGKRV
 101 YGLTVMIL CSVASGLSFG HEAKGVMTL CEFREWLFG IGGDYPISAT
 151 IMSEYANKT RGAFIAVFA MCGVGILAG FVALAVSSIF DKRPAPTYA
 201 VNRLSTPPQ VDYIIRIIV FGLPALTY YWRMKMPETA RYALVAKNI
 251 KQATADMSKV LQTDIELEER VEDDVKDPKQ NYGLFSKFL RRGLHLGT
 301 TSTWFLDLIA FYSONLFQKD IFSAGWIPK AATMNAHEV FRIARAOTLI
 351 ALCTVPGYW FTVAFTDTIG REKIDNGEF MMTVFMFAIA FPNHWIKPE
 401 NRIGFVWYS LTFEFANFGP NATTFIVPAE IFPARLRSTC HGISAAGKA
 451 GATGAGGFL YAAQSDKAK VDAGYPPGIG VKNSLIMLV LNTGLMFTF
 501 LVPEPKGSL EELSGEAEVS HDEK
 11AA_SEQUENCE 1.0
 ID 048640 PRELIMINARY; PRT; 524 AA.
 AC 048640;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE INORGANIC PHOSPHATE TRANSPORTER.
 GN PRT2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA:
 RA Mitsuoka N., Okumura S., Shibata D.;
 RL Soli Sci. Plant Nutrition 43:971-974(1997).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA: Pubmed-10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 features of the regions of 3,076,755 bp covered by sixty pl and TAC
 clones";
 RL DNA Res. 7:31-63(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: AB000094; BAA24282.1; -;
 DR EMBL: AB025638; BAA97416.1; -;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 DR Transmembrane.
 KW SEQUENCE 524 AA; 57644 MW; 54664DCD3ABED97 CRC64;
 Q48640 Length: 524 May 30, 2002 09:32 Type: P Check: 8027 ..
 1 MAEQDLGVK ALDVAKTQL HFTAIYIAGM GFPTDAYDLF CVSLVTKILG
 51 RIYFNPESA KPSLPPHVA AAVNGVALCG TISGOLFPGW LGRKLGKRV
 101 YGLTVMIL CSVASGLSFG NEAKGVMTL CEFREWLFG IGGDYPISAT
 151 IMSEYANKT RGAFIAVFA MCGVGILAG FVALAVSSIF DKRPAPTYA
 201 VNRLSTPPQ VDYIIRIIV FGLPALTY YWRMKMPETA RYALVAKNI
 251 KQATADMSKV LQTDIELEER VEDDVKDPKQ NYGLFSKFL RRGLHLGT
 301 TSTWFLDLIA FYSONLFQKD IFSAGWIPK AATMNAHEV FRIARAOTLI
 351 ALCTVPGYW FTVAFTDTIG REKIDNGEF MMTVFMFAIA FPNHWIKPE
 401 NRIGFVWYS LTFEFANFGP NATTFIVPAE IFPARLRSTC HGISAAGKA
 451 GATGAGGFL YAAQSDKAK VDAGYPPGIG VKNSLIVLV LNTGLMFTF
 501 LVPEPKGSL EELSGEAEVS HDEK
 11AA_SEQUENCE 1.0
 ID 048639 PRELIMINARY; PRT; 521 AA.
 AC 048639;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE INORGANIC PHOSPHATE TRANSPORTER.
 GN PRT3.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA:
 RA Mitsuoka N., Okumura S., Shibata D.;
 RL Soli Sci. Plant Nutrition 43:971-974(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA:
 RX MEDLINE-20181125; Pubmed-10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 features of the regions of 3,076,755 bp covered by sixty pl and TAC
 clones";
 RL DNA Res. 7:31-63(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: AB000094; BAA24281.1; -;
 DR EMBL: AB025638; BAA97415.1; -;
 DR EMBL: AB025638; BAA97415.1; -;

DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.
 SQ SEQUENCE 521 AA; 57257 MW; 5EB3FC6D42590E2F CRC64;

048639 Length: 521 May 30, 2002 09:32 Type: P Check: 4821 ..

1 MADQDLGLK ALDVAKTQLY HFAIVIAQM GFTDAYDLF CVSLVTKLLG
 51 RLTYNPTSA KPGSLPPHVA AAVNGVALCG TLAGQLFFGM LGDLGRKKV
 101 YGTLIMIL CSVAGSLSG NSAKGVMTTL CFFRWLGFG IGDYPLSAT
 151 IMSEYANKT RGAFTAAVFA MGVGILAGG FVALAVSIF DKRPSPYTE
 201 QDRFLSTPQ ADYIWRITVM FGLPALALTY YWRMKMPETA RYALVAKNI
 251 KQATADMSKV LQTDLEELER VEDDVKPPK NYGLFSKEFL RRHGLHLIGT
 301 TSTWELDLIA FYSQNLFOKD IFSALGMIK AATMNAHEV FKARAQTLI
 351 ALGSLVPGYW FTVAFIDILG RFAQLMGFF MMTVFMAFA FRYNHMLPD
 401 NRIGFVWYS LTFEFANFGR NATFTVPAP IEPARLSTC HGISAATGKA
 451 GAIVGAFGL YAAQPDQTK TDAGYPPGIG VKNSLIMLGV INFVGMLETF
 501 LVPEPKKSL EELSGAEVD K

11A_SEQUENCE 1.0
 ID Q9ZOX5; PRELIMINARY; PRT; 3449 AA.

AC Q9ZOX5;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, last sequene update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE PUTATIVE TRANSPORT PROTEIN.
 GN T10P11.5 OR AT4G02660.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OR NCBI_TaxID=3702;

RA SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RA Kaplan N., Johnson D., Schutz K., Gnoj L., Hoffman J., Tili S.,
 de la Bastide M., Granat S., Hameed A., Gottesman T., Hasegawa A.,
 Shohdy N., Parnell L., Dedhia N., Johnson A.F., Lodhi M.,
 Martienssen R., Chen E.Y., Wilson R., McCombie W.R.,
 "Sequence of A. thaliana BAC T10P11 from chromosome IV.";
 RT Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
 RL [2]

RA SEQUENCE FROM N.A.
 RA Spiegel L.A., Huang E.N., Nascimento L.O., de la Bastide M., Vil D.M.,
 Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
 Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
 Meves H.W., Lemcke K., Mayer K.F.X.;
 RA Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]

RA SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 3 WD REPEATS (TRP-ASP DOMAINS).
 DR EMBL; AC002330; AAC78268.1; -.
 DR EMBL; AL101495; CAB7751.1; -.
 DR InterPro: IPR000409; Belge_BEACH.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF02138; Beach. 1.
 DR Pfam: PF00400; WD40. 3.
 DR ProDom: PD007848; Belge_BEACH. 1.
 DR SMART; SM00320; WD40. 2.
 DR SMART; PS50197; BEACH. 1.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.

DR PROSITE: PS50082; WD_REPEATS_2; 1.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Hypothetical protein; Repeat; WD repeat.
 SQ SEQUENCE 3449 AA; 385006 MW; 6BB0DBA776690184 CRC64;

Q9ZOX5 Length: 3449 May 30, 2002 09:32 Type: P Check: 8548 ..

1 MKMGTLKDL KDKGVAETT ADLAGEAIS DPTPPSSSQ ASPSSFVAL
 51 AOHQNLNLSR TSRDKLIEL DFKRYWEEFR SSSSQEKKA ALNLSVNFEC
 101 RLVKQHAVND QLVTMLVEPH IFSFVIGRAF VADVEKLKDG SSHGANLFTA
 151 IEVLASGFPD KQSLDSCIL CCLIHFNAP LTVSASEGE KTVNVEEKE
 201 GSVVIMAKLA ASHPSAASL IEDDSIQLEF KVAAGSLMA FSRKVGIVS
 251 FHNQLHKNA MQLGLLVN DNGSTASYR KHLILKVLAM AVKFDPPDG
 301 DSATVTVGYD LLECEVELSY RPEYGVRLK DDIRNAGYH FLYQFALIS
 351 SMPKQIVPAF DHSPPHNRG SNDKQKPL SLKTRONDS EKQOSLSINS
 401 RQNDFFALKH FSPALSRLLD VLVTLAQTP IESSGTSYL LSQTKLTGYS
 451 RQTPSANNR YDEPCEGSG KVKDLAVQM LQDIFLKAEN KDLAEVLNR
 501 MKFTSHLE NYRICOELKT VPLLVLNMG FSSIQLEIL KILEYATVY
 551 NCVPEQELLS LCFLLQPID SELKHTILSF FVKLASFPDQ YKVLGEVGV
 601 LEVLQDLKQ HKLLRGPQY SGVSNHLDRV PPSPFKQHL DSODAIISP
 651 KLMESGSLR PIPEVERIT VGWDCMISL KSNQVQZAF RSANGVYIL
 701 PPLADEHRT SILRIFSLI TGDIKOVHNE ELEALIDVAK SGWATRVSGD
 751 QKLIYEVC DIMGALWRIY GVNGSAQRYF GEATGFSLL TTHTFQGE
 801 EQRDSHLMY YIKLFKHLR LITAVCENA INRMKHSVI TSQTFYDLY
 851 ESGLLCVLDE RHVQLLLEL ALEVLVPPFL TSEMSAEM AECEKASFLV
 901 KTAGQFND KQIYNAGAV RVLRSLLC TPRLQLEFLN LLERLARASP
 951 FKKEFLTSG CVELLETTY PRLQSSPFL SHALKIVEVL GAYRLSPBL
 1001 KMLCRYMOM RVNNGSPSLI GMMERLIME EDTGLECVSL APFVEMDSK
 1051 TGHASVQVSL GERSWPPAAG YSFVCWQVFR NFLTQELBS EYKAGGSSK
 1101 TPILSGQOSE QNIFRIPSVN AINSGPSYA ELYPQEDGIL TLATNSNSSL
 1151 SFSGLETEBG KWHHLAVHS KPNALAGLFQ ASVAVYIDG KLRHNGKIGY
 1201 SSPSPKSLD VVIGSATCA RACGGDSMAI LLDLPTDMS GIOKPEDSNR
 1251 QGDSKARCSG IYWDLDRLGN LSIQLPEKKL ITAFDCTGSE FMRATGSSL
 1301 VNLVPLSAA ASLIGGIPRF GRLVGNVSLC RQNVIGNSIR PVGMAVYLA
 1351 LVEAASRDM LHMALSLAC ALHONSQNVK DMETTYGYL LALFLRPKMA
 1401 LFDQCLEIF FOISACEAF SEPKKLESQ TTISMSPTI IPENNYEDPT
 1451 LCKFOYETSS VSGHGDMDF SGRKDSFSL SELEMGDNV ETSNCIVLSN
 1501 ADMVEHVLLD WTLWTAQVS IQIASLGFE NLISLWYS HNLAILRQIN
 1551 LVKHLVTLQ RGDVEVLVE KLVYLLRCL ENGFLTEPEL DVVPAITGF
 1601 NPPEIKONS SMRESMKHV IYRNLVLEML IDLQVYIKAE ELBQWHTV

```

1651 SSKLITYFLD GAVHPSSMRV IMTLGVCLT SSPNSLKEF ASGGYQGLVR
1701 VLOGEYDUSD IYYILFCLIF GKPVPRLPE VRMLDFHALM PDDGSHVELN
1751 FVDLDLSVVA MAKSTFDRLI MÖSMLAHÖSG NLÖOVARSVC AELVEGYADM
1801 TGELOGRAML HKTYAARLMG GEASAPATAT SVIRFMVLA KMPÖPSAAC
1851 KMEFLÖKCA DLYEVCVRAF HAVKLAKÖLS MKAEBÖNTG GDDSWEGNF
1901 CRVSHÖDMST KTSISAGSFP ÖDÖTSSVISTV DMVTPSDVVA VDKVENFLTT
1951 PPGESNSQFQ GREYIAKÖDG DHVGSVSASS EMKSLDLTGS SSQVÖPDSR
2001 SSESFSMALES PILSEKSSLE VFPIPSPSKS STISTPHPSH ISVEEDASS
2051 DOSSGSGÖSS AVHTLEFTIS KVLLETDSG YGGGPCSAGA SAVIDFMAEV
2101 CADIMTEÖIK AVÖALESILE MÖPLVYDEPC VVVFÖGICLS RVANNYLERRE
2151 LRÖDEEDDK LÖKRKSANL DAFÖMMIYDR VYMGAFPOPT GVLRTLEFLL
2201 SIÖLÖNKDG RVEEYSSGK GLLSIGRATR ÖLDAYVHSIL KNTNRTILYC
2251 FLPSFLITIG EEDLPSRLGI LVESTKQÖTS KLSGKESGID VSAVLÖLVA
2301 NKNITLCPSN LÖTDLNCCLC VNLISLHDÖ RKNÖVÖNÖASN IIKYLLVHRK
2351 SALEDLÖVKK PHRÖKÖFDVL HÖGFDRLITG NLPEFSKMLE SSÖÖITTKVL
2401 EÖGÖAVVMIÖ YIÖGSAKFPD VÖMKÖMDÖRR TREMGRLÖRD TSKLÖDKHME
2451 ÖVÖRRÖALE VVRDÖNSÄEL RYVRÖNKYGL ILHESVWPT HÖÖLÖVHERG
2501 IÖPWRISHGV EDLÖWÖLCPI EÖPYRMRKL ERÖKLÖIDSL HNDLEKÖTEL
2551 GEIÖLLKSKS EDGÖLVSDMD SEPÄFLTSEL YSEFSSEAD DÖKDVPÖSARN
2601 GÖVNDRÖTST NÄÄSLHNSLS FÖGKSSSTAV SVPIÖSVTDE KSEÖGSPÖKS
2651 SSGMÖDEIKH VEESEKELK DÖGEYLRPY LEHLEKIRFR YNCERYVGLD
2701 KHÖGIFLIGE LÖLYVIEÖNFY IDHÖGICÖEK ECEDELÖIID ÖAQÖLKKÖFH
2751 GSLESKSSKS TÖMSTTIKIG AVGÖRÖMAYG GÖGÖGKÖKVR VYÖNLPHPW
2801 MÖKÖDSVHEI LÖRÖYELÖRV AVEIÖFSDÖC NDLLVÖFKKE REEYFRNLÖLA
2851 MNLPRNSMLD TÖISSAKÖE SKESGRLPKL MÄKSFTKRMÖ NGEISNÖYL
2901 MHLNTLAGRG YSDÖLYÖVF PHLADYDGE SLÖLSDPNNF RÖLÖDKPMGCO
2951 TPÖGEEEFK RYESWDDPEV PÖFHYSHS SÄGIVLEFYLI RÖPFSÄENÖ
3001 KÖGÖGÖFDHÄ DRLFNSIRET WLSÄGKÖNT SDVÖELIPEF FYMPEFLENR
3051 FNDÖLGKÖKS GÖKVGDIYP PÖÄRSVREF IRKREÖLES DYÖSEMLHH
3101 IDLIFGKÖR GÖÄENAVNV FYHYTEGNV DVÖAVTDPÄM KÄSILÄQÖINH
3151 FÖGÖTPKÖLFQ KPHVKRRTDR KVPÖHLKHS MHLVPRIRK CÖSSINÖIT
3201 FNLKLLTGA NCLÖKPRGYK KYIRMGFPDR TLRFMSYDÖD KLLSTENLH
3251 EÖGÖIÖCAGV SHDGRIVTG AEDGLVSÖVR VSKDÖPGRSR RÖLÖKSÖLCA
3301 HÖAKVÖICLÖR SÖPYMMIÖSS SDDCTVILMD LSSLFVÖRÖL PÖNSVÖVTV
3351 YINDLGEIV TÄAGSVLÖWV SINGÖCLSV NTSÖLPÖDLI VÖVÄGÖTFSD
3401 WLETTVYVGT HÖSGÖLKVÖR MVHCTDPIKÖ CYRÖKÖKÖÖSE TESRRÖDN

```

11AA_SEQUENCE 1.0

```

ID 092WT3 PRELIMINARY; PRT; 516 AA.
AC 092WT3;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
OS INORGANIC PHOSPHATE TRANSPORTER.
OC Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=99087485; PubMed=9872450;
RA Okumura S., Mitsuoka N., Shitano Y., Shibata D.;
RT "phosphate transporter gene family of Arabidopsis thaliana.";
RL DNA Res. 5:261-269(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: AB005746; BAA34390.1; -.
DR EMBL: AB025638; BAA97413.1; -.
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR Prosite: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW Transmembrane.
SQ
SEQUENCE 516 AA: 56246 MW: 97828FENR894A145 CRC64;
Q92WT3 Length: 516 May 30, 2002 09:32 Type: P Check: 9638 ..
1 MANEQGSIL KALDVAKTQM YHVTAVVVG MGFTDSYDL FVLSLITKLL
51 GRIVQVPGS SSPGSLPDEI SNAVSGVAFV GTFIGQIFPG CLGDKIGRR
101 VYGLTLLIMT ICSICGSLG GRDPKTVMT LCPRFWMGF GIGGDYPLSA
151 TIMSEYSNKR TRGAFYAAYF GMÖGIGILAA GAVSLVSAY FESKFPBRAY
201 ILDGAATVP QADYVRIIL MVGALPALLT YVMÖKMPET ARTTALYSKN
251 AEOALDMTK VINVDIEASA AKNDQARVSS DEGLFPMKF LRRHGLHLG
301 TASTWFLDI AFYSONLFÖK DIFTIGWLP SAKTMNÄIÖE LYVIAKÖTI
351 IACÖSTVPGY FTVYGRIDM GRKKIÖMGF AMÄTIFMLSL ÄIYHNWTLP
401 ANRIGEVVLY STEFESNFG PNÄTFIVPA EIPPARIST CHÖISAASG
451 AGÄWGSFGE SALVKALGMS NLYIMAGIN LGLLLEFTI PETNGKSLSE
501 LSGETEPEKI KEKIVV

```

11AA_SEQUENCE 1.0
ID 09SOYL PRELIMINARY; PRT; 558 AA.
AC 09SOYL;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE HYPOTHEICAL 63.2 KDA PROTEIN.
GN F3M14.22 OR F18K10.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
 Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
 Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome 3 BAC F13M14 genomic sequence.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
 Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC F18K10 genomic sequence.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC011560; AAF51388.1;
 DR EMBL; AC013428; AAF76351.1;
 DR InterPro: IPR003441; NAM.
 DR Pfam: PF02365; NAM; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 558 AA; 63244 MW; F0BFC9A928C4874 CRC64;
 09SOY1 Length: 558 May 30, 2002 09:32 Type: P Check: 9617 ..
 1 MGRSSTVSLA PGFRFHPTDE ELVRYLKRK ICNKFKDA ISVVDYKSE
 51 PWDLPRNSRL KSRDLEWYFF SMLDKKRYNG SKTNBATEMG YWKTGKDR
 101 ILNSKVVGM KKTLYVHKR APERGTNMY MHEYRLVDQ LDKTVHQA
 151 FVLGRIFQKS GSGPKNGEY GAPVEEWE EEDDTFTFPD QEDGSEDHV
 201 YVHMDIDQK SENVVYDAI PIPLNFIHG SSNNVETNYS DSIWYIQTG
 251 NYMDSGGYFE QPAASYEKDQ KPIIRDQDS LQNEGICGV QDKHSETLOS
 301 SDNFGTDTs CYNDFVESN YLIGFAFLDP NSNLENDGL YLENDLSST
 351 QQDGFDEYD LTFEDEFDP SOLQMNDYF FDQELFOEV ETKLEKEET
 401 SRSKHVEEK EKDEASCQK VDADATEFEP DDGSISENAD KYKPLKKAS
 451 HMLAIPAPL ANASSEPFDK AAIRLHAQS SSVVHTAGM ITTDSNMGM
 501 SYKNENIDL ILSGLVQGN TAPKSGNS AMAMLIWCF WLLLSVSK
 551 VSILVSSR
 11AA_SEQUENCE 1.0
 ID 09AU01 PRELIMINARY: PRT: 540 AA.
 AC 09AU01;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PHOSPHATE TRANSPORTER 1.
 GN PRT.
 OS Lupinus albus (white lupine).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
 OX NCBI_Taxid=3870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PROTEOID ROOT;
 RA Liu J., Unde-Stone C., Li A., Vance C.P., Allan D.L.;
 RT "Characterization of phosphate transporters with enhanced expression
 in proteoid roots of white lupin.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF305623; AAK38196.1;
 SQ SEQUENCE 540 AA; 59383 MW; 49BBE0B807105D35 CRC64;

09AU01 Length: 540 May 30, 2002 09:32 Type: P Check: 5248 ..
 1 MAREIHVLN ALDYAKTQWY HSLQIIAGM GFTTAYDLF CISLVTLLG
 51 RIYTHVDCG KPGTLPPNVS AAVNGVAFCG TLSGLFFGWM LGDMKGRKV
 101 YGMTLMIMVI CSISGLSFG HSPNSVLATL CFFRWLFGG IGDYPLSAT
 151 IMSEYNNKT RGSFLAIVA MCGFGILGGG IFATISLAF KARDARAYE
 201 VDPVSTVPQ ADYIMRIIM VQALPALTY YWRMKMPETA RYVALVAKNA
 251 OOAADMSKV LQVEIQSETN KEBAQKPSR GLFSKEPLRR HGLHLG7AS
 301 TWELLDIAFY SQNLFOKDI SAIGWIPPAK TKNALDEYR IARQTLIAL
 351 CSTVPGWFT VALIDRGRF AIQLMGFFPM TVFMFALAP YDHWTRDNR
 401 IGFVYISLT FEFANFGPA TTFVPAEIF PARFRSTCHG ISSASGKIGA
 451 IYGAFFLYL AONKDKSKTD AGTPAGIGYK NSLYLVGVN ILGCFPTFLV
 501 PEPNCKSLEE MSGENEDEEP TKEGSYNNNN NNNRTPVPY
 11AA_SEQUENCE 1.0
 ID 09AU00 PRELIMINARY: PRT: 539 AA.
 AC 09AU00;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PHOSPHATE TRANSPORTER 2.
 GN PRT.
 OS Lupinus albus (white lupine).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
 OX NCBI_Taxid=3870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PROTEOID ROOT;
 RA Liu J., Unde-Stone C., Li A., Vance C.P., Allan D.L.;
 RT "Characterization of phosphate transporters with enhanced expression
 in proteoid roots of white lupin.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF305624; AAK38197.1;
 SQ SEQUENCE 539 AA; 59135 MW; 36C30F31F34742A6 CRC64;
 09AU00 Length: 539 May 30, 2002 09:32 Type: P Check: 478 ..
 1 MARDELSTLT ALDYAKTQWY HFLAIVAGM GFTTAYDLF SIANYTLLG
 51 RIYTHGAP KPGTLPPNVS LAVNGVALCG TLTGOVFFGWM LGDMKGRKV
 101 YGLTIAIMVF ASLASGLSFG HTAKGVSTL CFFRWLFGG IGDYPLSAT
 151 IMSEYNNKT RGAFLAIVA MCGFGILFGG IYSLYISITF DHASARAYE
 201 VDAEASLVPQ ADYIMRIIM LGALPALTY YSRMKMPETA RYVALVARNA
 251 KOAADMASKV LQVEIEAEO MVDKLSERN NNFGLFTRKF AKRKHILIG
 301 TCTTWFLDI AYSSNLFQK DIYSSIGWLP PAKENNAIHE VFRSKAOVL
 351 IALGTVPGY WTYVAFIDHL GREFIOLMG FFMVFMFAL AIPYDHTTK
 401 DNRIGELMV ALTEFFANFG PNSTFVVA BIFPARLST CHGISAAGK
 451 AGAIVGARG LYASQSKAT KRDYGPAGI GKKNTLIYLA VCNOLGNFFT
 501 FLVPSKSKS LBELSGENED EGTNETSQOP AASNTFVY
 11AA_SEQUENCE 1.0

ID 09AVR0 PRELIMINARY; PRT; 536 AA.
 AC 09AVR0;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PHOSPHATE TRANSPORTER.
 GN Ptl.
 OS Sesbania rostrata.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Robiniaceae; Sesbania.
 OC NCBI_TaxID=3895;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ROOT;
 RA Aono T., Oyaizu H.;
 RT "The expression of phosphate transporters in response to phosphate starvation is different from that of H⁺-ATPase or PP2C in Sesbania rostrata."
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AJ286743; CAC28218.1; -
 SQ SEQUENCE 536 AA; 58625 MW; CCDC7E1584277363 CRC64;

09AVR0 Length: 536 May 30, 2002 09:32 Type: P Check: 6827 ..

1 MAGELGVLNA LDVAKTQWYH FTAIVAGMG FETDAVDLFC ISLVTKLGR
 51 IYSDPAPAK PGLTPPNVQA AVTGVALCGT LAGQLFFGWL GDRLGRKKVY
 101 GITLLMWVC SLASGLSFGS SAKGVNATLC FFRFLGFCGI GDYPLSATI
 151 MSEANKKTR GAFIAVAFAM QCGFIMAGCI VALIVSAGFD HKKRVSYOE
 201 NPOASLVLPQ FDYWRIVLM FGALPALTY YWRKMEETA RYALVAKNA
 251 QQATDMKV LKVELEVEE KVQKTEADT NKYGLFSKEF AKRHGLHVG
 301 TTTTWFLDI AFYSQNLPOK DIFSAIGWIP PAKEMNAIHE VYIAAQTLL
 351 IALCSTVPGY WFTVAFIDYV GFALQLMGF FEMTVEMFAL AIPYDHWTKR
 401 ENRIGFVVMY SLTEFFANFG PNATTFVVA EIFPARLRST CHGISAAAGK
 451 AGAIVGAFG IYAAQSKDPT KTDKGYPTGI GIKNSLIMLG VINPAGMLPT
 501 LVPESKSGS LEBLSEGENE GAEGIELPAS ARTVPV
 11AA_SEQUENCE 1.0
 ID 09AVQ9 PRELIMINARY; PRT; 540 AA.
 AC 09AVQ9;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PHOSPHATE TRANSPORTER.
 GN Ptl.
 OS Sesbania rostrata.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Robiniaceae; Sesbania.
 OC NCBI_TaxID=3895;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ROOT;
 RA Aono T., Oyaizu H.;
 RT "The expression of phosphate transporters in response to phosphate starvation is different from that of H⁺-ATPase or PP2C in Sesbania rostrata."
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AJ286744; CAC28219.1; -
 SQ SEQUENCE 540 AA; 58947 MW; 7BA0EA8649BF187A CRC64;

09AVQ9 Length: 540 May 30, 2002 09:32 Type: P Check: 3365 ..

1 MAGGGGLGV LNALDVAKTQ WHFTAIIVIA GMSFTTAYD LFCISLVTKL
 51 LGRIYYTDYN KPRKGYLPPN VQAAVTGYAL CGTLAGOLFF GMDLGRKK
 101 KYGITLIMM VVCSIASGLS FGSSAKGWYA TLCPFRFNLG FGIGSDYPLS
 151 ATIMSEYANK KTRGAFITAAV FAMQFGIMA GCIYALVSS APDHKKYVPA
 201 YQEDAKSMV LPAFDYVMRI ILMFQAVPAA LTYWRKMP ETARYTALVA
 251 KNAQAASDM SKVLQVEIEA EEDKVQHLE SQNTQFLGS KEPAKRGHGH
 301 LVGTTTWFL LDIAFYSQNL FOKDIFTAIG WIPAKEMNA IHEVYIARA
 351 QTLIALCSTV PGYFTVAFI DYMGFPAQL MGFEFTVFM FALAIPYHHW
 401 TLKENRIGFV VMSLFFFA NFGPATFV VPAIFPARL RSCCHISAA
 451 AGKAGAIYVA FCFLYAAQST DPAKTDAQY TGICVKNLSI MGVINFFGI
 501 IFTLLVPSK GKSLEELSGE TEEDVEVAIE AAASSRTVPV

11AA_SEQUENCE 1.0
 ID 096243 PRELIMINARY; PRT; 524 AA.

AC 096243;
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PHOSPHATE TRANSPORTER.
 GN APTL
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ROOT;
 RX MEDLINE=9717798; PubMed=9025304;
 RA Smith F.W., Ealing P.M., Dong B., Delhaize E.;
 RT "The cloning of two Arabidopsis genes belonging to a phosphate transporter family."
 RT plant J. 11:83-92(1997).
 CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC EMBL: Y07681; CAA68945.1; -
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.
 SQ SEQUENCE 524 AA; 57610 MW; 62BD256C59634B5D CRC64;

096243 Length: 524 May 30, 2002 09:32 Type: P Check: 8141 ..

1 MAEQQLGVK ALDVAKTQLY HFTAVIAGM GFETDAVDLF CVALYKRLLG
 51 RIYFNPESA KQSLPPHYA AAVNGVALCG TLSQLFFGW LGDKLGRKKV
 101 YGLTLIMIL CSVASGLSFG NEAKGVMTTL CFFRFLGFG IGGDYPLSAT
 151 IMSEYANKT RGAFFIAVFA MGVGILAGG FVALAVSIF DKFPPAPYA
 201 VNRLSTPPQ VDYIRIIVM FGALPALTY YWRKMEETA RYALVAKNI
 251 KOATADMSKV LQTDIELER VEDVKDPRQ NYGLFSKEFL RRHGLHLGT
 301 TSWFLLDIA FYSQNLFOKD IFSALGWIPK AATNANATHEV FRARAKOTLI
 351 ALCSTVPGYV FYVAFIDTIG RKKIQLNGF MMTVEMFAIA FPNHMKPE
 401 NRIGFVVMYS LTEFFANLPG NATTFIVPAE IFPARLRSTC HGISAAGKA
 451 GATIGAFGL YAAQNDKAK VDACYPPGIG VKNSLIYLVG LNTIGMLPTF

501 LVPEPKGKSL EELSGEAHSV HDEK

11AA_SEQUENCE 1.0
ID 096264 PRELIMINARY; PRT; 524 AA.
AC 096264;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN APT2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RX MEDLINE=97177798; PubMed=9025304;
RA Smith F.W., Ealing P.M., Dong B., Delhaize E.;
RT "The cloning of two Arabidopsis genes belonging to a phosphate transporter family".
RL Plant J. 11:83-92(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: Y07682; CAA6946.1; -
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 524 AA; 57642 MW; CC12C41887127BCB CRC64;

Q96264 Length: 524 May 30, 2002 09:32 Type: P Check: 8941 ..

1 MAEQQLGLK ALDYAKTQLY HFTAIYAGM GFTDAYDLF CVDLYTKLIG

51 RIYFNPEPSA KPGSLPPHVA AAVNGVALCG TLGGLFFGWL LGDKLGRKKV
101 YGLTLVMNLT CSVASGLSFG HEAKGVMTTL CFFRFWLFGF IGCDDPLSAT
151 IMSEYANKTR KGAFIAVFA MGVGVLGAG FVALVSSIF DKRPAPTYA
201 VNRALSTPPQ VDYIWRITVM FGLPALATY YWRMKMPETA RYTLVAKNI
251 KQATADMSKV LQTDIELEER VEDDVKPRQ NYGLFSKEFL RRGHLHLIGT
301 TSTWFLDLIA FYSONLFQKD IFSALGIMPK AATMNATHEV FRIRAOGLI
351 ALCTSTVPGYW FYVAFTDTIG RKRIQLNGFF MWTVMFAIA FPNVHWIKPE
401 NRIGFVWYS LTFEFANFGP NATTFIVPAE IFPARLRSTC HGISAAGKA
451 GAIYAGFGL YAAOSODKPK VDAGYRPGIG VKNSLIMGV LNFGLMETF
501 LVPEPKGKSL EELSGEAHSV HDEK

11AA_SEQUENCE 1.0
ID 024029 PRELIMINARY; PRT; 538 AA.
AC 024029;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN LEPT1.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MONEYMAKER; TISSUE=ROOT HAIR;
RX MEDLINE=96408214; PubMed=9737001;
RA Darim P., Brunner S., Perisson B.L., Amrhein N., Bucher M.;

RT "Functional analysis and cell-specific expression of a phosphate transporter from tomato."
RT Planta 206:225-233(1998).
RL PLANTA 206:225-233(1998).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: Y14214; CAA74607.1; -
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 538 AA; 59042 MW; DB1ABE74C3272EF CRC64;

Q24029 Length: 538 May 30, 2002 09:32 Type: P Check: 4193 ..

1 MANDQVINA LDVAKTQLYH FTAIVYAGM FFTDAYDLFC ISMTYKLLGR

51 LYHHDGALK PGLSPRVSA AVNGVAFCGT LAGLFFGWL GDRKGRKKV
101 GWTLIMWVIC SIASGLSFGH TPKGVMITLC FFRFWLFGCI GGDYPLSATI
151 MSEYANKTR KGAFIAVFAV MGVGVLGAG VALIVSAFK GAPPAVEY
201 DAIGSTVQA DFWRTILMF GAIPACILTY WRMKMPETAR YTLVAKNLK
251 QAANDMSKVL QVELIABEPK VTAISEAKGA NDFGLFTEEF LRRHGLHLIG
301 TASTWFLDLI AFYSONLFQK DIFSAIGWIP PAQTWNALEE YKTIARAQTL
351 IALCTSTVGY WFTYAFIDKI GREAIQLMGF FMTYFMRL ALPHHWITLK
401 DHRIGFVWY SEFFEFANFG PNATTFVPA EIFPARLRST CHGISAAGK
451 AGAWGARGF LYAAQPTDPT KTDAGYRPGH MCEELVDRPW LCNFLGLMFT
501 FLVPEPNKRS LEDRENEG EETVAEIRA TSGRTVPV

11AA_SEQUENCE 1.0
ID 004660 PRELIMINARY; PRT; 938 AA.
AC 004660;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE A_TM021B04.3 PROTEIN.
GN A_TM021B04.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Dante M., Wamsley P., Gibson A.;
RT "The sequence of A. thaliana TM021B04."
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Wash-U;
RT "The A. thaliana Genome Sequencing Project."
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Waterston R.;
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF007271; AAB61068.1; -
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR001320; Ion_glu_receptor.
DR InterPro: IPR001311; SBP_glu_receptor.
DR Pfam: PF01094; ANF_receptor; 1.
DR Pfam: PF00060; 11g_chan; 1.
DR SMART: SM00079; PBPc; 1.
SQ SEQUENCE 938 AA; 105169 MW; 7831B9C30E96CE12 CRC64;

```

004660 Length: 938 May 30, 2002 09:32 Type: P Check: 5621 ..
1 MKRNNLVLS LLEFVIVFLM QVGEAQNRT NVNAGIVNDI GTAYSNMNTLL
51 CINMSLSDFY SSHPETQTRL VTTVDSKND VYTAAGTR IPTSLNHLAL
101 RYCVINHOF ICTLLIYIA FSSKALDLIT NKEVKALGP WSMQAOQPMI
151 EMQKSOVPI VTYSATSPSL ASIRSQYFPR ATYDSSOVH AIKEIIRLKG
201 WREYAPYVD DTFEGIMPR LTDVLQELNV RIFRYTISP NATDDEISVE
251 LLMKMLPTR VEVVHVELL ASRFAKATE IGLMKQCYW ILTNITTDVL
301 SINTEIET MOGVIGVKT VPSKELNF RSRWTKRFPPI SDLNVYGLMA
351 YDATTALALA IEAGTSNLT FVKMDAKRNV SELQGLGVSQ YGPKLLQTL
401 RVRFQGLAGD FOFINGELQ SFVEIVNNG QGGRTIGFM KEYGLKKNVD
451 QKPASKTFPS SWODRLRPII WPGDTTSVPK GWEIPIINGKR LQIGVEVNTT
501 FQOFVKATRD PITNSTIFS FSIDYEA VI QAIPIYDISD FIPEQDGYD
551 ALYYQVYLGK YDAVADTTI SSNSMIVDF SLPTPGCVG LVPVYDSVR
601 RSSITFLMPL TLALWLISL SEFIIGLVW VLEHRVNDP DGPQYQLST
651 IFWFSEFIMV FAPREVLFS WARVVIIWY FLVVLVLOSY TABLASLTT
701 QHLHPTNTI NSLAKGESV GYQSSFIIGR LRDSGEFAS LVSYSPEHC
751 DALISGQAE GGVSAVIMV PYVRIELQY CNKYKWQTP FKVDGLGFVR
801 PIGSPYVADI SRALKVEES NKANOLENAM FKPIDESCPD PLTNPPNPS
851 VSRQGLGDFS FVVLFLYVAI VCTALLKLV YQFLKEPNQ RNLRVLMKF
901 NEPDQSYIK DVTKCOCSSG QGMPKNGEG ANAVNNGN

!!AA_SEQUENCE 1.0
ID 022055 PRELIMINARY; PRT: 542 AA.
AC 022055;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
GN PT1.
OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; eusterids I; Gentianales; Apocynaceae; Rauvolfioideae;
OC Vincaceae; Catharanthus.
OX NCBI_TaxID=4058;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. PRETTY IN ROSE;
RA Kai M., Masuda Y., Kikuchi Y., Osaki M., Tadano T.;
RT "Isolation and characterization of a cDNA from Catharanthus roseus
RT which is highly homologous with phosphate transporter.";
RL Soil sci. Plant Nutrition 43:227-235(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AB004809; BAA20522.1; -.
DR InterPro; IPR003662; sub-transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 542 AA; 59580 MW; B7B761A255C08BEDA CRC64;

022055 Length: 542 May 30, 2002 09:32 Type: P Check: 2442 ..
1 MAKEQLOVLN ALDVAKTQWY HFTAVINGM GFTDADLF CISTVYKLLG

```

```

51 RIYYHDGAV KPGLTPPNVS AAVNGVAFCG TLAGQLEFGW LGDKMKRKY
101 YGMTLMIMVI ASIAGLSFG DKPRVAMTL CEFREWLFGG IGGDYPLSAT
151 IMSEYANKT KGAFLAANFA MGGFGLAGG MVAITVASF KAGFPAPAQ
201 DGAVASTYPE ADYWRILIM FGAIPALTY YWRKMBETA RYALVAKNA
251 KQANDMSKV LQVELAEQK KVEKPAQBP NAFLGFTKEF LKRHGLLLG
301 TATWFLDDI AFSQULFOK DIFSAIGWP PAQTMNIEE VEFIAAQLT
351 IALCSTVPGY WFTVELIDRI GFETIQMGF FEMTWFMAL AIPYNMTHK
401 DNRIGEVIWY SLTFEPANFG PNATTFVVA EIPPARLST CHGISAAAG
451 AGALIGAFGF LYAAPSDFS KTDKGYPPGI GYKALLVYG CVNPLGMVFT
501 FLVPEAKGKS LEEVSKNEE EVENGTELRQ QSGHDTPTVP VL

!!AA_SEQUENCE 1.0
ID 09ST22 PRELIMINARY; PRT: 536 AA.
AC 09ST22;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; eusterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. XANTHI-NC;
RA Kai M., Adachi H.;
RT "Phosphate transporter.";
RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AB020061; BAA86070.1; -.
DR InterPro; IPR003662; sub-transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 536 AA; 58909 MW; C2AA6E8E040504F CRC64;

09ST22 Length: 536 May 30, 2002 09:32 Type: P Check: 2750 ..
1 MAKDLOVLNA LDVAKTQLYH FTAIVIAMG FETDAYLFC ISLVTKLGR
51 IYYHRDGLK PGLSPNVSA AVNGVAFQGT LAGQLFPGWL GDKMGRKYY
101 GMTLMNVIC SIASGLSPFH TPKGVMITLC FFRWLOFGI GGGYPLSATI
151 MSEYANKTR GAFIAAVFAM QGFGILAGM VAIIVSAFK GAFAPPTYHA
201 DALASTVSOA DEYWRILIM GAIPAAPTY WRMKMBETAR YTALVAKNIK
251 QANNDMSKV QVEIEEOKK VENVSQNTGN EFGFSEFEL RRGHLLLT
301 ASWFLDLIA FYSQNLFOK IFSALGWIP AETWNALEEV YRIARQTLI
351 ALGSTVPGW FTVFIDRIG RFAIDLMGF FMTVFEALA IPYHWTLDK
401 NRIGEVIMYS LLEFFANFGP NATTFVPAE IIPARLSTC HGSMAAGNA
451 GAMIGAFGL YAAOPTDPK VDAGYPAGIG VRNSLIVGC VNELGVETP
501 LVPEKSKSL EEMSRNEGE EESGTEKNS GRTVPV

!!AA_SEQUENCE 1.0
ID 09FH47 PRELIMINARY; PRT: 1112 AA.

```

AC 09FH47;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE GB|AD55616.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=20181125; PubMed=10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
 RT clones.";
 RL DNA Res. 7:31-63(2000).
 DR EMBL: AB022212; BAB08882.1; -
 DR InterPro: IPR000822; Znf-C2H2.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 SO SEQUENCE 1112 AA; 12709 MM; 7C1A579B826DE85 CRC64;

09FH47 Length: 1112 May 30, 2002 09:32 Type: P Check: 6231 ..

1 MASSFKKQVM AERLYKDAED LIAKGNHKA LEILEDLISI HMNANDALL
 51 HIEGQNFVD LAEKEDRDM EVAYLLGVA CFSQHYVFE LCAOSLEMLG
 101 RNLSVSHYK KCLKAREAL SDSNSKLS STPVAFRAIV SRKKDLEGE
 151 IEDASWIAA SKTSPLOKPE PKRYDEKTP EPSKDFORF RSYWGLDYK
 201 IKRDELAKSI AKLTSEYGI RKRKQDALE QALASKDKK WTFMCRKTC
 251 SKRCSSEVC RKHEQOQDA DFKSSEKDI VKRIGNMWAR KISLGMEPV
 301 DAVAAVEIK NQLADYKAF SSKNGMSYE WPLAVDEERK QLLQEIKLIL
 351 VSFCLAKLS GSRDWYMHF PVYTLAKLEY SKOGLIDSHL SETPOSICFL
 401 ECHELTRILD FLKTRICKRN DGTQVCTAY DSVLGRIOYK VKIDFPOFS
 451 FLUDKRLILK INDVQPDDE GTINVDPSA HYAKAPVHGD DIIMLIDYN
 501 SVDKTEPPPI REHNLDIWLA VLKAVQFPCR TLGNRYAKKV QVVDYDAALT
 551 DVENMCVSEN ERKRLPEDQ WSRVYSLDC VCEERYPKNS LTRKLEVRAY
 601 RDVFGALHP TLDFLDEDC LNFTRHKSJL SDDKVLQALD LKSVVQKY
 651 LLMOTKILLI DNRISILINN LTRLSAFDNR TYIILQLKFE LLMELVNMES
 701 KAKSDAABAD LNLLEKEKL QPREKQSKR KPOSKRRDR TKRPSTIS
 751 SLIDKVEHK PESTPSLRT VEEDSMERP ALASETDMS RHGDEPLSDH
 801 LESAGEEAA RYNSALDMTL KALSTIKIF EDLMONSOI QDQROEYPT
 851 ALONLEFAY SEALQDEGVY SCLLSDFLS QEFILSMSSD AAKVVAVALN
 901 LMRCKNPER ESLVTRFTL VENKRMSCRK CRTITNSPVQ SSYGVAMAD
 951 STRELKAFG NIKFVDILKL IRLGYKLC NKKGCGKKS YVDHTRCP
 1001 PIETIVLEME KSETEKEISE TTRALDWEID ISRLYEGLLE TINTNRLVSM
 1051 VCGGGEERH VCLAYEKNRM VNLRRCEGAG EDVGMKNVY RFGGRKVRP
 1101 ELIYESVRL MA

11AA_SEQUENCE 1.0
 ID 09AYT3 PRELIMINARY; PRT; 537 AA.
 AC 09AYT3;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PHOSPHATE TRANSPORTER.
 GN NTP2.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; easterids I; Solanales; Solanaceae; Nicotiana.
 NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kal M.;
 RT "phosphate transporter.";
 RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AB042950; BAB21545.1; -
 DR SEQUENCE 537 AA; 59246 MM; DE2D0D68B6C22361 CRC64;

09AYT3 Length: 537 May 30, 2002 09:32 Type: P Check: 5587 ..

1 MAKDQLOVLN ALDYAKTOLY HFTAIYAGM GFTDAYDLF CISLVTKLIG
 51 RIYHHGAP KPGTLPRNVS AAVNGVAFG TLAGQLFEGM LGDKMKRRV
 101 YGMTLMMYI CSIASGLSFG HTPKSVWTL CFFRFWLGFG IGGDYPLSAT
 151 INSEYANKRT KGAFIAVFA MGFEGILLAG MVAIIYSAF KGAPPAOTYQ
 201 TDLGISTISO ADFWRITLM FGALPAMTY YRMKMPETA RYTLALVAKNL
 251 KQANDMSKV LOVDIEEOE KVENVSQNTN NEGLFSKEF LRRICHLIG
 301 TASTWFLDI AFYSQNLQK DIFSAIGWP PAOTNALDEE YKTIARAQTL
 351 IALCSTVRGY WFTVEFIDKI GREFAQLMGF FMTYEMFAL AIPYHHTLK
 401 DNRIGFVIM SLTFEPANFG PNATTFVYPA ELPFARLST CHGISAAAGK
 451 AGAMIGAFGE LYAAQPTDRK KADAGYPAGI GVRNSLIYLG CVNPLGNVET
 501 FLVPSKGS LEEMSRNEG EESGTEKN SGRTPVY
 11AA_SEQUENCE 1.0
 ID 09AYT2 PRELIMINARY; PRT; 523 AA.
 AC 09AYT2;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PHOSPHATE TRANSPORTER.
 GN NTP3.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; easterids I; Solanales; Solanaceae; Nicotiana.
 NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kal M.;
 RT "phosphate transporter.";
 RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AB042951; BAB21562.1; -
 DR SEQUENCE 523 AA; 57206 MM; EDC878CDB3CFF91A CRC64;

09AYT2 Length: 523 May 30, 2002 09:32 Type: P Check: 7879 ..

1 MSADNNLOVL NALDYAKTOL YHFTTIVYAG MGFPTDAYDL FSLVTKLL
 51 GLIYTKSDL LKPGTLPPYV SASVGYVALY GTLAGQLFEG WLGDKMGKRT
 101 YGMTLIIMY LCSVASGLSF GSTPKGYMAT LCFRFWLGF GIGGDYPLSA

```

151 TIMSEYANK TRGAFIAAVF AMQGFILFS GIILIVSAG FDHAYVPSF
201 QENALSTVP QADYIWRIL MEGALPAFLT YWWMKMPET ARYTALVAKD
251 AKRAAQOMGR VLOVEIDPED AKIEQMSRDE TNKFGLSWE FVGRHGLHLF
301 GTCTSWLDD IAFYSQNLQ KVFESAVGMI PAATMNAVQ EYKIAAQT
351 LIALSTVPG YWFTVAFIDI VGRFAIQLMG FEFMTVEMFA IAIPEHMTK
401 KDNHIGFVAM YALTFEFANF GPNATTFVVP AEIPPARLRS TCHGISAAG
451 KAGAIYVAG FLYAASKDP TKTDAQYPAG IGINSLLAL GAINALGMVC
501 TFCVPESKGR SLEASQETI TEE

11AA_SEQUENCE 1.0
ID 065383 PRELIMINARY; PRT: 523 AA.
AC 065383;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DE PHOSPHATE TRANSPORTER.
GN NPPT4.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RA SEQUENCE FROM N.A.
RT "Phosphate transporter."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB042956; BAB21563.1;
SQ SEQUENCE 523 AA; 57242 MW; 9F4C9B6E73CF7DE2 CRC64;

09AT1 Length: 523 May 30, 2002 09:32 Type: P Check: 9177
..
1 MSADNNLQVL NALDVAKTQL YHFTIYIAG MGFTDAYDL FSISLYTKL
51 GLUYTKREL LKPGTLPTV SASVGVAVL GTLAGQLFEG WLDGDKGRKT
101 VYGTTLIMV LCSVASGLSF GSTPKGVMT LCFRFLNGF GIGGDYPLSA
151 TIMSEYANK TRGAFIAAVF AMQGFILFS GIILIVSAG FDHAYVPSF
201 QENALSTVP QADYIWRIL MEGALPAFLT YWWMKMPET ARYTALVAKD
251 AKRAAQOMGR VLOVEIDPED AKIEQMSRDK TNKFGLSWE FVRRHGLHLF
301 GTTSTWELDD IAFYSQNLQ KDVSAGIMI PPAKTMNAVQ EYKIAAQT
351 LIALSTVPG YWFTVAFIDI VGRFAIQLMG FEFMTVEMFA IAIPEHMTK
401 KDNHIGFVAM YALTFEFANF GPNATTFVVP AEIPPARLRS TCHGISAAG
451 KAGAIYVAG FLYAASKDP TKTDAQYPAG IGINSLLAL GAINALGMVC
501 TFCVPESKGR SLEASQETI TEE

11AA_SEQUENCE 1.0
ID 065383 PRELIMINARY; PRT: 783 AA.
AC 065383;
DT 01-AUG-1998 (TEMBLrel. 07, Created)
DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)
DE 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
GN F12F1.17 PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

```

```

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RA SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA.
RA Vysotskaya V.S., Osborne B.I., Schwartz J.R., Tortum M., Yu G.,
RA Kuan A., Oji O., Liu S., Buehler E., Conway A.B., Conway A.R.,
RA Dewar K., Feng J., Kim C., Kurtz D., Palm C.J., Li Y., Shinn P.,
RA Sun H., Davis R.W., Ecker J.R., Federpspiel N.A., Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC F12F1 sequence, complete
RT sequence."
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC002131; AAC17615.1;
DR InterPro: IPR003864; DUF221.
DR Pfam: PF02714; DUF221.
SQ SEQUENCE 783 AA; 89528 MW; 2AE3FA091DE5BF5A CRC64;

065383 Length: 783 May 30, 2002 09:32 Type: P Check: 9544
..
1 MATIGDIGVA AAINILTAII FLAPAILRI QPFNDRYFP KWLKGISS
51 PLHSGALVSK FVNVNLSYL RELNMPPAL KMPPELIDH AGDSAVYLR
101 IYLGLKIFV PIALLMMSIL VPVNMTSHGL OLAKLRNVS SDIDKLISIN
151 IENGSDSLYF GREWHLVMA YAFETWTCYV LMKEYEKVAA MRLAFLONEQ
201 RRPQFTNLG LSQLSQVLY RVNVPADPDS ISDSVEHFEL VNHDPHYLTH
251 QVYVNAADLA ALVEQKKSTQ NMLDYQOLKY TRNQHPRRI KTFGLGMCK
301 KVDAIDHYIA EIEKLNQER KVKKDDTSV MPAAVSFKT RMGAASAOQ
351 QQSSDPTEWL TEWAPEAREY EWSNLAIPIY SLTVRLIMH IAFELTFEFP
401 MIPAFVQSL ASIEGIEKNA PLKSLIEND LKSVIGGL PGIVLKLFLI
451 FLPSILMWS KKEGFVSLSS LERRAARFY IENLINFLG SVTGSAPFQ
501 LDSELKQSAK EIPKTVGVAI PIKATFEITY IMVGMGAGIA GEILRLKPLI
551 FFHKNGLVL KTEKDREEM NGQINYNAT EPRQLVELL GLVYAVTPV
601 ILPFIITFFA LALVVRHOI INVYNOYES AARWPDVHG RIISALITAO
651 ILMGLLSTK GAAQSTPFL FLPIITFEFH RYCKGRYEP FLRHPLKEAM
701 VKDTLERARE PNENLKPYLO KAVIHVPKD NDYEDSRFDE ISGYCIEDSD
751 EECYVPTKR QSRINTPAVS HASRGSRSRP PSK

11AA_SEQUENCE 1.0
ID 022148 PRELIMINARY; PRT: 1374 AA.
AC 022148;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DE 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE PUTATIVE NON-LTR RETROELEMENT REVERSE TRANSCRIPTASE.
GN AT2G45230.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RA SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Buell C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umeyan L., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,

```


RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant *Arabidopsis*
 RT *thaliana*.";
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA Lln X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: ACC02387; AAB82639.1; -;
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR002156; RnaseH.
 DR InterPro: IPR000477; RVTse.
 DR InterPro: IPR000130; Zn_MTpeptidase.
 DR Pfam: PF00075; rnaaseH; 1.
 DR Pfam: PF00078; rvt; 1.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 KW RNA-directed DNA polymerase.
 KW SEQUENCE 1374 AA; 158442 MW; 718BA300AE61F581 CRC64;
 SQ

022148 Length: 1374 May 30, 2002 09:32 Type: P Check: 8740 ..

1 MRLISWNCQG VGNTPTVYRHL REIRGLYPEE VIFLCETKKR RNYLENVGH
 51 LGFDLHIVE PIKSGGLAL MKKDSVOIKY LQSDRLIDA LLIMQDEEY
 101 LTCIYGEVQ AERGELMERL TRLGLSRSP MMLTDFDEL VDPEKIGCP
 151 ARKSSCLEF RQMLNSCGLW EYVNSGYOPS WYGNRNDLY QCRLDRTVAN
 201 QAMMELFPOA KATYLOKICS DISPLINLY GDMNRKMGF KYDKRWYQRE
 251 GFKLLCWFV SQOSTKTNAL MMEKIASCR EISKKKRYSK PSSAVRIQEL
 301 QFKLDAATKO IPPDRRELAR LKKEISOEYN NEEOFWQES RIMMRNGDR
 351 NTKYFHATK NRRQNRIOK LIDEGRENT SDEDIGRAVE AYFKLFASE
 401 DVGTVLELE NLTPVSDQM NNMLAPITK EYVQATPSI NPHICPPDG
 451 MNGELYQOFW ETMGDDITEM VOAFRRSGSI BEGMNKTNIC LIPKILAEK
 501 MTERPISLC NVLYKVIKGL MANRLKIIP SLISETQAAF VKGRILSDNI
 551 LIAHELHAL SSNNKCEEF IAKTDISKA YDREMPFLE KAMGLGFAD
 601 HWIRLIMECV KSVRYOVLIN GTPHGEIIPS RGLROGDPIS PYLFVICTEM
 651 LYKMLQSAEQ KNOITGLKVA RGNRPISHL FADDSMFYCK VNDEALGOII
 701 RIIEEYSLAS GQRYNYLAKS IYFGKHISE RRCLYKRRKLG IERGGEGYV
 751 LGLPESFOGS KVAITLSTLKD RLGGKVLGWO SNELSPGKE ILLKAVAMAL
 801 PLYTMSCEFI PKTICOQIES VAAEFWMKNK KEGRGLHWKA WCHLSRPKAY
 851 GGLGFKEIEA FNIALGKOL WMITPEKDSL MAKVFKSRVF SKSDPLNAPL
 901 GSRPSFANKS IYEAQVLIKO GIRAVIGCE TINVWTDPMI GAKPAKAQA
 951 VKRSHLVSOY AANSIHVVKD LLLPDGRDN WNLVSLFPD NTQENILALR
 1001 PGKETRHRF TWEXSRSGHY SVKSGYWMY EIIINORNPQ EVLOPSIDPI
 1051 FOQIWKLDVP PKIHFLMRC VNNCLSVASN LAYRHLAREK SCVCPESHGE
 1101 TVNHLLFCFP FARLTWALSP LPAPGGEWA ESLFRNMHHV LSVKSPDEE
 1151 SDHALLIWI LMRILMKNRND LVFKGREFTA PQVILAKED MDAMNNKEP
 1201 QPQVTSSTRD RCYKWDPPSH GWKCNCTDGA WSKDLGNCVY GWVLRNHTGR

1251 LIMLGLRALP SQOSVLETEV EALRMWAVLSL SHFNRYRYF ESDSQYIVSL
 1301 IONEMDIPSL APRQODIRNL LRHFEVYKFO FTRREGNNVA DRTARESLSL
 1351 MNYDPKWTYSI TPDWIKNLVD LETV

11AA_SEQUENCE 1.0
 ID 082193 PRELIMINARY; PRT; 550 AA.

DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE ATG19880 PROTEIN.
 GN ATG19880.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA MEDLINE=20083487; Pubmed=10617197;
 RA Lln X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant *Arabidopsis*
 RT *thaliana*.";
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA Lln X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: ACC005169; AAC62128.1; -;
 DR SEQUENCE 550 AA; 62247 MW; 951AA77DAB02AE11 CRC64;
 SQ

082193 Length: 550 May 30, 2002 09:32 Type: P Check: 6852 ..

1 MSTLSDISAI LFSLSRAFTS PFAVFOYQO CIGCLLALG WLALEYVRNR
 51 EYKRINKNIK AGNSLAFYQ DINELHSRQ VKLPRVSYVM PLKFGGEHNL
 101 HNMRSQITSL YGGRLEFLV VESTEDPAVN AVSRLLSMYQ DHYEAKVVA
 151 GLSTCSQKI HNOIGVERM HKOTKYVLEL DDDVALRHGT IGALTTEMEK
 201 NPEVCSMGF ATGRTFFLM GCGMMHADD PRODRGYVS GLRGGYSDD
 251 MTLASLAGAH KRLTSPDVA VEPHPLASDL SEGRWYMYLR KQTVLESYI
 301 SKVMNMINKA LEAVHYCLSM GFVAPYVMAI IHITSALRIY IKGYHOLEDT
 351 TSASGGLSYF LMLHIDRRM FYVGMVLVIT LAICFTIELL SMMVLTRREV
 401 QLCNNLSDEA PRSLATYVM GLPFIIGHI SPWYRGFNKP RTGVFWNVED
 451 YPIPLVYHT IFQIKSALE NKHFVETPL LLQTVSSSIW LSTSLFRCI
 501 LIKESFTSSA ASMPETMYK GLNGKTSYLV FARAQNIHI SNIMASRMC

11AA_SEQUENCE 1.0
 ID 082362 PRELIMINARY; PRT; 787 AA.
 AC 082362;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE ATG46060 PROTEIN.

GN AT2G46060.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 RN NCBL_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eissen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC005397; AAC62897.1; -
 DR InterPro: IPR000361; EGF-like.
 DR PROSITE: PS00022; EGF_1; UNKNOW_1.
 SQ SEQUENCE 787 AA; 87537 MW; B6E497738D1D542B CRC64;
 OS262 Length: 787 May 30, 2002 09:32 Type: P Check: 3197 ..
 1 MATELFLLR TLVVFSCALS ICNCLQDAD NNGNFTVSS FRYPESEVRP
 51 YDRYIRVLDL PMWFSSLVNA IESVDITAK SISKISKL PYICFRDGP
 101 PLPDASTNAL OGELGRFEN GSFERAQDSE IAQCQYPMOK NITRLRLNEQ
 151 CLVYLILIST NIDTAQIVRS SAFSSANIS VEGCKTITMW GPCPCNQTTP
 201 LSCSRFDNQT ASVISCADSE PSSCLTGAET KYVALDVGI AEOLVIMASN
 251 VKVDSNESYL MCVARPEAVA SETLHDVAAD IHKVPILVNK PKAGRYIYI
 301 SLGGRKREFA QGTNSSSRVC FSINKYVLCG PYKAGGNCG QQIYIIQAVM
 351 RRGMLRPFQS YFEPVNDASL GSSSTNPELE PIVSNFSSIP ELDTSTWTF
 401 LMNIPQSGG GHIFRLSLD STIOEYVLR FGGLPTIDR DYYVVRTSA
 451 SRMFSFLYN SSKEMVDFYI LYAREGTSF GLROLIDSNT PAASRGSPTL
 501 VSLSLERCPR GCSYQGCYR AFDANGLTSY RFFIYLESK ICKHNDNCS
 551 CDRTHGGFDC SIEIVSHOEI IVQSIALIAS NAAALPAYW ALKQREYEW
 601 VLTSSGSSS ALYHACDVGI MCVLSYVNLQ FMEFWLSFMA VVGTFYVLT
 651 AGEAVKRTIH TVVAITLALL ALTOATRASN IIVLAIGSL GLLIGLVEF
 701 VTKYRSYCS AGFSLMLDR PRAVKEWFSN LIKTLKKRFR WGFVAGVIA
 751 FTMMAISFKI AIVNHNLAH NGADNYELTR QDSLSRN
 11AA SEQUENCE 1.0
 ID 092076 PRELIMINARY; PRT; 131 AA.
 AC 092076;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE AT2G03400. PROTEIN.
 GN Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 RN NCBL_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eissen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC006284; AAD17436.1; -
 DR EMBL: AC006284; AAD17436.1; -
 SQ SEQUENCE 131 AA; 14837 MW; 095DCE6344F868D3 CRC64;
 OS2076 Length: 131 May 30, 2002 09:32 Type: P Check: 6617 ..
 1 MVOQSSTL TICGSYKVS LLRNRLNSVK ASSLIGRCV SCQFLKSPS
 51 FRSHKSLKQ RNLRLVEARN PQGGGEGCL DPSSRESEA NEDILFFRO
 101 LDLATRVQVK TKCRFISFK MVANGFDVLQ Y
 11AA SEQUENCE 1.0
 ID 09FRL1 PRELIMINARY; PRT; 654 AA.
 AC 09FRL1;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE HYPOTHETICAL 73.0 KDA PROTEIN.
 DE F22H5.13.
 GN Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 RN NCBL_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
 RA Malt R., Ronning C.M., Koo H., Fujii C.Y., Uteback T.R.,
 RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC025814; AAG12693.1; -
 DR HSSP: P16435; IBIC.
 DR InterPro: IPR001094; Flavdn-1-like.
 DR InterPro: IPR001226; Flavodoxin.
 DR Pfam: PF00258; flavodoxin; 1.
 DR PRINTS: PR00369; FLAVODOXIN.
 KW Hypothetical protein.
 SQ SEQUENCE 654 AA; 72997 MW; BADD77B16845516D CRC64;
 OSFRL1 Length: 654 May 30, 2002 09:32 Type: P Check: 7509 ..
 1 MSTTSSVRV LAFVALSAT TFYCIHKYRR LKHKNLSLN PSSTLKASG
 51 KIFFISQGT AALALQRLHE LCASNDIAFD IVDPHSYEPD DLPKELVLAF
 101 IASTWDGKP PKNGEFLVMN LGESAEPRV GSLLSDCKF AVGVGSRAY
 151 GESYNAVAKL LSSRMIGLGG LEMIPVSGD VDDGELDRAP QDWCGVIVY

```

11AA SEQUENCE 1.0
ID 022301 PRELIMINARY; PRT; 537 AA.
AC 022301;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE PHOSPHATE TRANSPORTER.
GN MPT2.
OS Medicago truncatula (Barrel medic).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_TaxID=3880;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98086876; PubMed=9425684;
RA Liu H., Trieu A.T., Blaylock L.A., Harrison M.J.;
RT "Cloning and characterization of two phosphate transporters from
RT Medicago truncatula roots: regulation in response to phosphate and to
RT colonization by arbuscular mycorrhizal (AM) fungi.";
RL Mol. Plant Microbe Interact. 11:14-22(1998).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: AF000354; AAB81346.1; -
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 537 AA; 58738 MW; B3BD94029D8CB8C0 CRC64;

022301 Length: 537 May 30, 2002 09:32 Type: P Check: 33 ..

1 MSGELGVINA LDVAKTQLYH FTTIYIAGMG FFTDAYDLFC ISLVTKILGR
51 IYTERPNPTR PGTLPSPSAOS AVTGVALVGT LAGQLFEGWL GDXLGRKKVY
101 GLTLLIYVVC SVAGSLFSGS SPKSYMATLC FFRWLFGCI GGDYPLSATI
151 MSEVANKTR GATIAVFAM QGFGLGGI VALIVASIFD HKKVPFEE
201 NPASTSLVPQ FDIYVRLIIM FGALPAALTY YWRKMPETA RYALYAKNA
251 KQAADMSKV LOVELEVEEE KYOKMTSDR NSYGLFSKOF AARHGLALFG
301 TCSWPLLDI AFYSQNLFOK DIFSAIGWIP PAKEMNAIHE VYKIAAQTL
351 IALCSTVPGY WFTVAFIDHM GRFAIQMGF FFWTVFMEGL AIPTYDMSKE
401 ENRIGFVVMY SLTFFSNFG PNAATFVPA EIFPARLST CHGISAAAGK
451 AGAIVGAFGF LVAAQSKDPT KTDKGYPGI GIKNSLIMLG VINFGMLCT
501 LVPESKSGS LEELSGENEG EGAEATEDEG PRFENVA

```

```

11AA SEQUENCE 1.0
ID 022302 PRELIMINARY; PRT; 533 AA.
AC 022302;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE PHOSPHATE TRANSPORTER.
GN MPT2.
OS Medicago truncatula (Barrel medic).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_TaxID=3880;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98086876; PubMed=9425684;
RA Liu H., Trieu A.T., Blaylock L.A., Harrison M.J.;
RT "Cloning and characterization of two phosphate transporters from
RT Medicago truncatula roots: regulation in response to phosphate and to
RT colonization by arbuscular mycorrhizal (AM) fungi.";
RL Mol. Plant Microbe Interact. 11:14-22(1998).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: AF000355; AAB81347.1; -
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 533 AA; 58250 MW; 3C4AB00747E4135C CRC64;

022302 Length: 533 May 30, 2002 09:32 Type: P Check: 2173 ..

1 MSGELGVINA LDVAKTQLYH FTTIYIAGMG FFTDAYDLFC ISLVTKILGR
51 IYTERPNPTR PGTLPSPSAOS AVTGVALVGT LAGQLFEGWL GDXLGRKKVY
101 GLTLLIYVVC SVAGSLFSGS SPKSYMATLC FFRWLFGCI GGDYPLSATI
151 MSEVANKTR GATIAVFAM QGFGLGGI VALIVASIFD HKKVPFEE
201 NPASTSLVPQ FDIYVRLIIM FGALPAALTY YWRKMPETA RYALYAKNA
251 KQAADMSKV LOVELEVEEE KYOKMTSDR NSYGLFSKOF AARHGLALFG
301 TCSWPLLDI AFYSQNLFOK DIFSAIGWIP PAKEMNAIHE VYKIAAQTL
351 IALCSTVPGY WFTVAFIDHM GRFAIQMGF FFWTVFMEGL AIPTYDMSKE
401 ENRIGFVVMY SLTFFSNFG PNAATFVPA EIFPARLST CHGISAAAGK
451 AGAIVGAFGF LVAAQSKDPT KTDKGYPGI GIKNSLIMLG VINFGMLCT
501 LVPESKSGS LEELSGENEG EGAEATEDEG SRV

```

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: AF022873; AAB82146.1; -
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.
 SO SEQUENCE 538 AA; 58699 MW; 9B1961F0CB7EC0A6 CRC64;

022548 Length: 538 May 30, 2002 09:32 Type: P Check: 4793 ..

1 MANDOLINA LDVAKTQLYH FTAIVAGMG FETDAYDLFC ISMVKTLGR
 51 LYHHGQALK PGLSPNNVSA AVNGVAFCGT LAGOLFEGWL GDKMGRKKYV
 101 GMTLMINVC STASGLSFGH TPKGVMITLC FFRWILGFGI GGDYPLSATI
 151 MSEYANKKTR GAFIAAVFAM QGFGILAGM VAIIVSAFK GAFPAAYEV
 201 DAIGSTVPOA DEVMRIIME GAIPAGLITY WRKMKPETAR YTAIVAKNLK
 251 QAANDMSKVL QVEIEAPEK VTAISEAKGA NDFGLFTKEF LRRHGLHLG
 301 TASTWELDI AFYSQNLFOK DIFSAIGWIP PAOTMNALEE VYKIRAOQL
 351 IALCSTVPGV WFTVAFIDKI GRFAQLMGF FPMVTFEAL ALPHYHWTIK
 401 DHRIGVVMV STEFFPANGF PNATTEVPA EIPPARLST CHGISAAAK
 451 AGAMVAGFGF LYAAOPTDPT KTDAGYPPGI GVRNSLIVLG CVNGLMLFT
 501 FLVPSNGKS LEDLSRENGC EEFYVAELIRA TSGRTVYV

11AA SEQUENCE 1.0 PRELIMINARY; PRT; 528 AA.

ID 022549
 AC 022549;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE INORGANIC PHOSPHATE TRANSPORTER.
 GN LEPT2.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OC NCBI_TaxID=4081;
 OX RN
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liu C., Muchhal U.S., Mukatira U., Kononowicz A.K., Raghothama K.G.;
 RT "Tomato phosphate transporter genes are differentially regulated in
 RT plant tissues by phosphorus";
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC EMBL: AF022874; AAB82147.1; -
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 DR Transmembrane.
 KW SEQUENCE 528 AA; 57762 MW; 38165A56E2DE488C CRC64;

022549 Length: 528 May 30, 2002 09:32 Type: P Check: 3242 ..

1 MANGDNDNNN LQVNALDLA KTQLYHFAI VIAGMGFTD AYDPSISLV
 51 TKILGRLYYT KPDLLKPGT PPAVSASVYG VALVGLAGQ LFGWLGDKK
 101 GRKKYGMVL VLMVCSVAS GLSFGSTPKG VMTTLCEFRF WLGFGIGDY
 151 PLSATIMSEI ANKKTRGAFI AAVPAGQFCG ILFSGIYALI TAAGDHAYR
 201 SPTEENAL STVPSDYIM RIIMFGSLP AALTYWRMK MPETARYAL
 251 VAKDARRAQ DMGKVLQVEI ESEAKIEQI SRNETNQFGL FSEMFRRHG

301 LHLFGTCSTW FLIDIAFYQ NLFOKDFSA VGMIPKAPYM NAVQEVYKIA
 351 RAOTLIALCS TVPGWFTVA FIDIGRFPI QLMGFFMTV FMAIAPYH
 401 HMTLEANRIG FIVMSLTFF FANFGPNNT FVPAELFPA RLRSIOHGS
 451 AAGKAGAIY GAVGLYIAO SKDPNKTDAG YPAGIGIKNS LYLGCINAL
 501 GMLTCEVPE PKGSLEAS OETITGEA

11AA SEQUENCE 1.0 PRELIMINARY; PRT; 396 AA.

ID 09XEL6
 AC 09XEL6;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE HIGH-AFFINITY PHOSPHATE TRANSPORTER P1 (FRAGMENT).
 OS Triticum aestivum (wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae; Poideae;
 OC Triticeae; Triticum.
 OC NCBI_TaxID=4565;
 OX RN
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=CV. CHINESE SPRING;
 RA Li Y.J., Chen S.Y., Li Z.S.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: AF110180; AAD26146.1; -
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 DR Transmembrane.
 KW NON_TER
 FT NON_TER
 FT SEQUENCE 396 AA; 43519 MW; A2ACD6A657E0B25E CRC64;

09XEL6 Length: 396 May 30, 2002 09:32 Type: P Check: 7287 ..

1 FETDAYDLFC ISLVTKLGR IYYREGADA PGLSPNNVSA AVNGVAFCGT
 51 LSQLDFEGWL GDRMGRKKRY GMTLMCWILC SIASGLSFGS TIGSVNATLC
 101 FFRFWLFCGI GGDYPLSATI MSEYANKKTR GGFIAVFAV QGFGIITGV
 151 VTLVSAFR AAPTOAYOD APLASTPAQA DEVMRIIME GAVPALMTY
 201 WRKMKPETAR YTAIVAKNLK QAADMSKVL QVDIGAEED PRANCYGA
 251 DDRNSFGLFS GEFLRRHGLH LLGTRATCWL LDIAFYSONL FOKDIFTAIN
 301 WPKRATMSA LEEVHIRARA QTLIALCGTV PGYFTVALI DRIGRFWQL
 351 GGFPAVAVM LDIAFYHNM TTPGNHIGFV VIALTFEFA NGCPNA

11AA SEQUENCE 1.0 PRELIMINARY; PRT; 382 AA.

ID 09ZPP4
 AC 09ZPP4;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE THH13.9 PROTEIN.
 GN THH13.9.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 OX RN
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=CV. COLUMBIA;
 RA Washu;
 RT "The A. thaliana Genome Sequencing Project.";

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA.
 RA Drone K., Nguyen C.;
 RT "The sequence of A. thaliana T3H13.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA.
 RA Waterston R.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: AF128396; AAD17374.1; -;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.
 SQ SEQUENCE 382 AA; 41929 MW; 979C03576443F8D7 CRC64;

09ZPP4 Length: 382 May 30, 2002 09:32 Type: P Check: 2196 ..

1 MYPHGGCSR RYGFPTDSYD LEVLSLITKL LGRITYQVPG SSSFGSLPDC
 51 ISVAVSGVAF AGFPLQIIF GCLGDKLGRK RYGLTLILIM TICSIASSL
 101 FGKPKTVMV TLCEFRWLG FGIGDYPIS ATIMEYANK RTGAFIASV
 151 FAMGVGTLA AGVSLVST LEIEFPSPA YILGASTV PODYVRII
 201 LMGALPALL TYYRMKME TARYTALVAK NAEQALDMN KEITVEDK
 251 RFALTCOSE NCLIDLGRY WIIQIFGMV TVFMLVLAVP YHMTLPANR
 301 IGFYIFST FFFSNGFPA TFIYPAEIEF PARISTCHG ISASGRAGA
 351 MVSFGEFAL GISLELTGE TPERVKEKI GV

11AA_SEQUENCE 1.0
 ID 09SMG9 PRELIMINARY; PRT; 540 AA.
 AC 09SMG9;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE INORGANIC PHOSPHATE TRANSPORTER.
 GN PTL.
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_Taxid=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NAMJAK;
 RA Yun S.J., Baek S.H., Lee M.C.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: AF156695; AAD38859.1; -;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.
 SQ SEQUENCE 540 AA; 59150 MW; 3593F4335A8C568 CRC64;

09SMG9 Length: 540 May 30, 2002 09:32 Type: P Check: 8707 ..

1 MANDQVLNA LDVAKTQLYH FTAIVYAGMG FPTDAYDLFC ISWYTKILGR
 51 IYHHNDNALK PGLSPNVSA AVNGVAFCT IAGOLFEGML GDKMKRKKV
 101 GMTLMINVC SINSGLSFH TRKSVTILC FFRFLGRCI GGDPLSNTI
 151 MSEYANKTR GAFIAVEAM QGFGLAGM VAIIVSSAFK GAFPAVEV

201 DALASTVSOA DFWRIIME GAIPAGLITY WRMKPETAR YTALVANKLK
 251 QANDMSKVL QVELEAPER VGAISEAKKA NEFGFSKEF LRRIGLILIG
 301 TASTWELLDI AFYSQNLFOK DIFSAIGWIP PAQTMNALEE YKIAKQTL
 351 IALCSTVPGY WFTVAFIDRI GREAIQLMGF FFMVTFEMAL ALPYHMTLK
 401 DNRIGFVMY SLTFFPANFG PNATTFVPA EIFPARLST CHGISAAGK
 451 AGAMGARGF LYAQPDPK KTDAGYPPGI GVRNSLIYLG CVNELMILFT
 501 FLVPSNGKS LEEMSRNEG EETVAEMBA TSGRTVLKFF

11AA_SEQUENCE 1.0
 ID 09LIS5 PRELIMINARY; PRT; 537 AA.
 AC 09LIS5;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE INORGANIC PHOSPHATE TRANSPORTER.
 GN PTL.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_Taxid=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yun S.J., Baek S.H., Lee M.C.;
 RT "Differential regulation of tobacco phosphate transporter genes by
 phosphorus.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF156696; AAF4025.1; -;
 DR EMBL: AF156696; AAF4025.1; -;
 SQ SEQUENCE 537 AA; 59198 MW; 75820D1C35C2CF01 CRC64;

09LIS5 Length: 537 May 30, 2002 09:32 Type: P Check: 6167 ..

1 MAKQDQVLN ALDVAKTQLY HFTAIYAGM GFPTDAYDLF CISTVTKLIG
 51 RIYHHDCAP KPGTLPPNVS AAVNGVAFEG TLAGOLFEGW LGDMKRRV
 101 YGMTLMNVI CSIASGLSFG HPEKSVMTL CFPRFWLFG IGGDYPISAT
 151 IMSEYANKKT RGAFIAVFA MGFGLAGG MVAIVSAF KGAPADTYO
 201 TDPLGISTVQ ADFYWRILIM FGAIPAMTY YRMKMPETA RYTALVAKNL
 251 KOANDMSKV LOYDIEEEOE KVENVSQNTN NEFGFSKEF LRRIGLILIG
 301 TASTWELLDI AFYSQNLFOK DIFSAIGWIP PAQTMNALEE YKIAKQTL
 351 IALCSTVPGY WFTVAFIDRI GREAIQLMGF FFMVTFEMAL ALPYHMTLK
 401 DNRIGFVMY SLTFFPANFG PNATTFVPA EIFPARLST CHGISAAGK
 451 AGAMGARGF LYAQPDPK KADAGYPAGI GVRNSLIYLG CVNELGAVFT
 501 FLVPSNGKS LEEMSRNEG EESGTETMKN SGRTPV

11AA_SEQUENCE 1.0
 ID 09MSKO PRELIMINARY; PRT; 265 AA.
 AC 09MSKO;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE PHOSPHATE TRANSPORTER (FRAGMENT).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_Taxid=4530;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, INDICA;
RA Yu F., Zhang A., Chen S.;
RT "Rice phosphate transporter.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF229169; AAF40188.1; -.
FT NON_TER 1 1
FT NON_TER 265 265
SQ SEQUENCE 265 AA; 28329 MW; DC88DD1D52E06A74 CRC64;

O9M5K0 Length: 265 May 30, 2002 09:32 Type: P Check: 8753 ..

1 FFTDAVDLFC ISLVSKLGR IYYTDIASDT PGSILPPVSA AVNGVALCGT
51 LAGQLEFGWL GDLGRKRSYV GFTLVLMVVC SVASGLSFGK TAKGVVATLC
101 FFFRWLGFGI GGDYPLSATI MSEYANKRTR GAFIAVFAVM OGFGILFGAI
151 VALVSNAGFR NAYPAPSYAD GRAASLVPEA DYYWRITLMF GTRPGGAQLL
201 LAHENARRTA RYTGSAQOTR SKAGRXHXQG VYTRXFXQGX XAPGKRGVXR
251 RXXPRXXXKG XFFXG

!!AA_SEQUENCE 1.0
ID O9M562 PRELIMINARY; PRT; 393 AA.
AC O9M562;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE PHOSPHATE TRANSPORTER (FRAGMENT).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, JAPONICA;
RA Yu F., Zhang A., Chen S.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF239619; AAF42956.2; -.
FT NON_TER 1 1
FT NON_TER 393 393
SQ SEQUENCE 393 AA; 42860 MW; E45E1E7A5A042B9B CRC64;

O9M562 Length: 393 May 30, 2002 09:32 Type: P Check: 15 ..

1 FFTDAVDLFC ISLVSKLGR IYYTDIASDT PGSILPPVSA AVNGVALCGT
51 LAGQLEFGWL GDLGRKRSYV GFTLVLMVVC SVASGLSFGK TAKGVVATLC
101 FFFRWLGFGI GGDYPLSATI MSEYANKRTR GAFIAVFAVM OGFGILFGAI
151 VALVSNAGFR NAYPAPSYAD GRAASLVPEA DYYWRITLMF GTRPGGAQLL
201 WRKKMPEEAR YTALARRNAK QAAADMSKVL DTEIQEDADR ABAVAAGGAG
251 NEWGLFSRHE VRRHGVHLVA TTSTWFLDI AFYSOILFOR DIFSKGWIP
301 PARTNAVEE VERIARAQAL IALCGTIPGY WFTVAFIDVA GREAIQIMLF
351 AMMTVMGLL AAPYHHWTP GNHTGFPVVM GTFEPFANFG PNA

!!AA_SEQUENCE 1.0
ID O9MAX4 PRELIMINARY; PRT; 709 AA.
AC O9MAX4;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE TRANSPORTASE.
OS Zea mays (Maize).

```

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPORTIN-1;
RA Yan X., Li J., Cowperthwaite M., Fu H., Dooner H.;
RT "Jittery, a low-copy, mu-related transposon apparently mobilized by
RT BSMV infection.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF247646; AAF66982.1; -.
SQ SEQUENCE 709 AA; 81941 MW; F1080D806D982ABF CRC64;

O9MAX4 Length: 709 May 30, 2002 09:32 Type: P Check: 7892 ..

1 MSHLEFSPWQ SSSSFVPCDE HSLKPVIGMS FDSIDELEGF YRTYAHEGCF
51 SVRIGAGKK NDVEHKRFV CSREGFTRC AEAKNQKHF ETRCGCMARV
101 YRILGODKRY YIASFVEEN HGLVSPDKIP FLNSNRTICQ RAKTTLFTCH
151 KASIGTSQAV RLQVSDGFD NIGCMKRDLO NYRGUNEKI KNADAQLFYA
201 QMERKKEANS AFFYFAVDE HGKLYITCWA DATCRKSYTH FGDLYSVDAT
251 YSTNOYMRP APTGVNHHM QVFEFGAFL ANEKIESYEW LFTPLVAMG
301 GKAPRLIITD EDASIKSAIR TTLPTIHLR CMHIMEKVS EYGHPTSHD
351 KEWDALNTC VNGSETPPEF EKRNNALMDA YGLESNEMLA NRYKIESNI
401 PAFFMDTILA GVLRTSRSE SANSEFNRI HRKLCFEVFW LRPDTALERO
451 RHBEIKADHI SIHSPVLRT PMWEKQASI LYTHKVKIF QEEVLIARDH
501 CSYLGITQOD AKFVYVSDG SMRDYVQWC TSNIFGSCC KLEFKGICIC
551 CHILLANGE KLYELSSYI LKRMETCRK ECVYDDGNT LBEKPDANE
601 AERKRKITIV RNKIEAIOR AKSSNEAMDF LVSSVLNIGE SLGHVSSNV
651 QPTQEEYENF ICGKIPADIQ IHPPDVNSK GRSKRIRRAK ELPSKSKGN
701 ARKDMCAPL

!!AA_SEQUENCE 1.0
ID O9IKL5 PRELIMINARY; PRT; 390 AA.
AC O9IKL5;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PHOSPHATE TRANSPORTER (FRAGMENT).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JINGXIL17, AND CV, JAPONICA; TISSUE=SHOOT;
RA Yu F., Zhang A., Chen S.;
RT "Rice shoot phosphate transporter.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF271893; AAF76345.1; -.
FT NON_TER 1 1
FT NON_TER 390 390
SQ SEQUENCE 390 AA; 43066 MW; 61A281BB286366E1 CRC64;

O9IKL5 Length: 390 May 30, 2002 09:32 Type: P Check: 9448 ..

1 FFTDAVDLFC ISLVKILGR IYYFNPAKS PGSILPPVSA AVNGVAFGCT

```

51 LAGOLFPGML GDKMRKKVY GMTLMVLVIC CLASGLSFGS SAKGVMTLIC
 101 FFRFMLGRGI GGOYPLSATI MSEYANKRTR GAFIAAVFAM GGFNLNGGI
 151 VAIIVSAFK SRDAPAYRD DRGTGVPPK TTRGASCSCS AP1PALLTY
 201 WRMKPETAR YTALVAKNAK QAAADMTQVL NVEIVEONK ADEVARROQF
 251 GLEFNOFLRR HGRHLGTV CMFVLIDAFY SSNLFOKDIY TAVOMLPKAD
 301 TMSALEEMFK ISRAQTLVAL CGTIPGWTFT VFFIDILGRF VIQLGFFTFM
 351 TAFMLGLAVP YHMTWTPGNH IGFVVMYAFY FFFAXXGPYA

11AA_SEQUENCE 1.0
 ID Q9ATE1 PRELIMINARY; PRT: 323 AA.
 AC Q9ATE1
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE PHOSPHATE TRANSPORTER JX17 (FRAGMENT).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 GN NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ming F., Shen D.,
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF335588; AAK25766.1; .
 DR InterPro: IPR002173; PFKB.
 DR PROSITE: PS00584; PFKB_KINASES_2; UNKNOWN_1.
 FT NON_TER 1 1
 FT 323 323
 FT NON_TER 1 1
 SO SEQUENCE 323 AA; 34038 MW; 17688B53225530F4 CRC64;

09ATE1 Length: 323 May 30, 2002 09:32 Type: P Check: 2612 ..

1 LECISLVTKL LGRITYRVDS SPSPGTLPPH VSASVNGVAF VGLTSGOLFF
 51 GWLGDKLGRK RYGITITLMIM VILCSLASALS FGHITPTSWA TLCEFFRWLG
 101 FGIGGDYPLS ATIMSEYANK KTRGAFIAAV FAMQGFGLIT GGLVAIIVSA
 151 SERAFAFAPR YGEDPVASTP PQADFWRITL IMLGALPAL TYWRITKME
 201 TARTALVAN NAKQAADMS KYIQVYEMRN IGNNGSGRRA VRAYLRRVCP
 251 AARAAPGCHV GDVVAAGCHV LQEPYVEGH IORGVDPOG GDERAGGAV
 301 PHRAGADADR AVRDCARLLV HCG

11AA_SEQUENCE 1.0
 ID Q9AXEO PRELIMINARY; PRT: 333 AA.
 AC Q9AXEO
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE PHOSPHATE TRANSPORTER (FRAGMENT).
 OS Oryza rufipogon (Wild Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 GN NCBI_TaxID=4529;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ming F., Shen D., Zhang Y.,
 RT "Common wild rice phosphate transporter mRNA."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF337531; AAK06857.1; .
 FT NON_TER 1 1
 FT 333 333
 FT NON_TER 1 1

SQ SEQUENCE 333 AA; 36461 MW; 9C95C37D45BF0EB0 CRC64;
 Q9AXEO Length: 333 May 30, 2002 09:32 Type: P Check: 5023 ..

1 LECISLVTKL LGRITYENPA SKSPGSLPPN VSAAVNGVAF CGTLAQOLFF
 51 GWLGDKMGKR KYVGMTITLMIM VICCLASGLS FGSSAGVNA TLCEFFRWLG
 101 FGIGGDYPLS ATIMSEYANK KTRGAFIAAV FAMQGFGLIT GGIVAIIVSA
 151 AKSRFDADA YRDDRAGSTV PQADYANRIV LMEGALPALL TYWRMKME
 201 TARTALYAK NDKRAADMA RVLNVELVDE OEKAAATATA AAEERARE
 251 QYGLFSREPA RRHGHHLGT TYCMEVLIDIA YYSQNLFOKD IFSALGWPE
 301 AKTMSALDEL YHARAQTLI ALGCTVPGYV FTV

11AA_SEQUENCE 1.0
 ID Q9MIT0 PRELIMINARY; PRT: 535 AA.
 AC Q9MIT0
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PHOSPHATE TRANSPORT PROTEIN.
 GN T5N23.60.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eustoidia II; Brassicales; Brassicaceae; Arabidopsis.
 GN NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Obermaier B., Ottenwaelder B., Duchemin D., Zeitler K., Mewes H.W.,
 RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.,
 RL Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL138650; CAB77590.1; .
 SO SEQUENCE 535 AA; 58332 MW; 8FA274A5413FF800 CRC64;

Q9MIT0 Length: 535 May 30, 2002 09:32 Type: P Check: 3263 ..

1 MAGDQNLVYN ALDVAKTQWY HTFAIITAGM GFTTDAYDLF CISLVTKILG
 51 RIYHVDSSE KPGTLPPNVS AAVNGVAFCG TLAGOLFFCG LGDKLGRKKV
 101 YGMTLMVAVL CSIASGLSFG SNPKVTWTL CEFREWLGRG IGGDYPLSAT
 151 IMSEYANKKI RGAFIAVFA MQGFGILTGG IFALIIVSAF EAKFPAPTYO
 201 IDALASTVPO ADYVWRITLM VGLPAAPTY YSRKMPETA RYALAVADA
 251 KLAASNMKV LQVEIEAQO GTEDKSNSFG LFSKEPMKH GLHLGTST
 301 WLLIDIAFYS QNLFOKDIFS AIGMIPPAQT MNAIQVEYKI ARAQTLILC
 351 STVPGTWFTV AFIDVIGREA IQMMGFFFTV VEMFALAIPT DHWTHKENRI
 401 GFVAMYSLTF PFANFGPNAT TFWVPAEIRP ARFRSTCHGI SAASGLGAM
 451 VQAFGEIVIA QSPDKTKEH GTPPGIGVKN SILVIGVNL LGMVFTLLVP
 501 ESKGKSLEEM SGENQDNDES SSSSNNNSNN AVSTA

11AA_SEQUENCE 1.0
 ID Q9LZ39 PRELIMINARY; PRT: 428 AA.
 AC Q9LZ39
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE HYPOTHETICAL 48.9 KDA PROTEIN.
 GN T22P11.220.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
 RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
 RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL162971; CAB86000.1; -
 KW Hypothetical protein.
 SO SEQUENCE 428 AA; 48883 MW; 425E34860CE1E18 CRC64;

Q9LZ39 Length: 428 May 30, 2002 09:32 Type: P Check: 3770 ..

1 MTLPFLAAV FVLQSLTL VAEIKSFYIS NDSRPVILL KFGIIEIGHV
 51 TVSVSVSVL SPIDSSKLG FVLSSESLP HVLLEQONF SFCVLDSHYI
 101 LHFTFPDLS PPPRSQFSKS YPTSPNDYS LFFANCVPET RVSMKYHTEI
 151 YHDLYPNGSR DYLLAGSAQL PGLYLVFPLC YLSFLCFWLC FCMNHRQIVK
 201 RIHLMTALL LVKSLTLICA AVYKHVRYT GTAHGWNIVF YIQFISVYL
 251 LEWIVYLGN GMSFLKPKIH VKKKLVIV VPLOYLANIA SIYIGEGPY
 301 TDQWVSNQI FFLADITCCC AIVFAMVSM CCLRETSKTD GKAVKMLAKL
 351 PVLRKFYVLV IGYLFFTRIV VYVMKRADE TYQVSNAAE EIATLSFYCL
 401 MFYMFRIEK NEYCDVDEE EIVELSLK

11AA_SEQUENCE 1.0
 ID Q9ASF0 PRELIMINARY; PRT; 794 AA.

AC Q9ASF0;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 85.9 KDA PROTEIN.
 GN P0456F08.9.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-CV. NIPPONBARE;
 RC Sasaki T., Matsumoto T., Yamamoto K.;
 RA "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, PAC
 RT clone:P0456F08.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AF002901; BAB39409.1; -
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; kinase; 1.
 DR SMART: SM00219; Tyrc; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Hypothetical protein; Serine/threonine-protein kinase;
 Transferase.

SO SEQUENCE 794 AA; 85935 MW; 289C9C2539F1A1AC CRC64;

Q9ASF0 Length: 794 May 30, 2002 09:32 Type: P Check: 5843 ..

1 MYDGLAIIP LPTSPVPVF PIRSYPLP ANKPYPNPA VSPASIHPT
 51 NHGKAHGVP IAAHKKERHH SMLVNTNHN THAGPYVAP KGRHHSLPV
 101 NNTRVKGPAY SPNSPSIHR KHGIPVAPR KOHSSNLPPS HHRPHKGSFP
 151 VISPTPKAD NASATKHGRS GLHSPAPAP VGLPPSEGNA RGNPAPAPRH
 201 PHEHSPSNS PELADIDIAS VLMKSOVRV MGAARATEDP EKVIVLIDLV
 251 PLGEKFPDXT ALLVFERFW KOVNOFTSSV CVCVSVSLP SPPTAPRMNN
 301 GLSNVNDPRL HPLAVDGNH RETKSGILV IIVLSVAF ILSGALVI
 351 CFKIRNRNHL TEESPMPPK AGPGSAVNGS RLGSRPISAS PSFSSSLVTV
 401 KGTAKTESLI EMERATQRPD NSRIIEGGF GRVYEGILED GERVAVKILK
 451 RDDQVYREF LALEMLSLR HHRNLVKLG ICTEHIRCL VVELVPNGSV
 501 ESHLHGSDBG TAPLDMDARL KIALGAARL AYLEDSSPR VIHRDFRSSN
 551 ILLEHDETPK VSDGLARTP IEGNEHIST RWMGTGGYVA PEYAMTGHL
 601 VKSDVSYGV VILELITGRK PVDILRPQO ENLVAMACP LTRSDGLETI
 651 IDPLSGSIL FDSIAVVAI ASMCVQPEVD QRPNGEVQ ALALYCDGGS
 701 EFNESGSPSO DLHIQDSGI SRASLDVVE PVSAELFNA SAHYDTLDAS
 751 GSFRRYSSSG PLRVGRTGHN RERGLSTGSS SEHGTOFR IDSE

11AA_SEQUENCE 1.0
 ID Q9ARI9 PRELIMINARY; PRT; 543 AA.

AC Q9ARI9;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PHOSPHATE TRANSPORTER 1.
 GN P11.
 OS Lupinus albus (White lupine).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteeae; Lupinus.
 OX NCBI_TaxID=3870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liu J., Uhde-Stone C., Li A., Vance C.P., Allan D.L.;
 RT "A phosphate transporter with enhanced expression in proteoid roots of
 white lupin (Lupinus albus L.)."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY026321; AA01938.1; -
 SO SEQUENCE 543 AA; 59687 MW; D754FC2C7DA515F4 CRC64;

Q9ARI9 Length: 543 May 30, 2002 09:32 Type: P Check: 1485 ..

1 MARHIVLN ALDVARTQWV HFTAIINGM GEFIDAVDLF CILVYKRLGS
 51 RIYHYVDGAG KRGTLPPNVS AAVNGVAPCG TLSQGLFFGW LGDKMKRKY
 101 YGWTIMTAVI CSIGSLSPG HSPNSVLATL CFFRFWLGFG IGGDYPISAT
 151 IMSEYSNKTI RGSFIAVFA MGFEGILGG IFALIIAFAF KAFEDAPAYE
 201 VDPYGVTPQ ADYIMRIYV VQALPALTY YWRMKMDETA RYALVAKNA
 251 QQAAMDSKV LOVEIQSETN KEAQGRPSF GLFSKEPLRR HGLHLGTAG
 301 TWFLDIAFY SONLFOKIDF SAIGWIPAK TMNLDEYR IARAQTLIAL

351 CSTVPGWTF VALIDRGRF AIQLMGFFPM TVFMALAI P YDHWT HKDNR
 401 IGVVIXISLT FFFANGPNA TTFVPAEIF PARFRSTCHG ISSASRKLGA
 451 IYAGFGFLYL AONKDKSKTD AGYPAGIVK NSLIVIGVN IIGCFCTFLV
 501 PBNKSLKEE MSGENEDEEP TRECSYNNNN NNNNNNRTV PYV

11AA_SEQUENCE 1.0
 ID Q40115 PRELIMINARY; PRT; 214 AA.
 AC Q40115;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE RIBONUCLEASE (RNASE Lc1) PRECURSOR.
 OS Luffa cylindrica (Smooth loofah) (Sponge gourd).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Luffa.
 OX NCBI_TaxID=3670;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nakamura T., Sasaki K., Funatsu G.;
 RT "Cloning and sequencing of two cDNA clones coding for ribonucleases
 from the seeds of sponge gourd (Luffa cylindrica).";
 RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: D64011; BAA10891.1;
 DR InterPro: IPR001568; RNase_T2.
 DR Pfam: PF00445; Ribonuclease_T2; 1.
 DR PROSITE: PS00531; RNASE_T2_2; UNKNOWN_1.
 KW signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 214 RIBONUCLEASE (RNASE Lc1);
 FT SEQUENCE 214 AA; 24080 MW; 0F50D18D5D358BE CRC64;
 SO Q40115 Length: 214 May 30, 2002 09:32 Type: P Check: 3380 ..

1 MAMAKREIVL VEVLTILFPM VKSQTFDSFW MVQMPPAVC SFOGRVCVQ
 51 GLSFTTHGV WPQKGTSTVI NCGPTFPDT KISHLESTLN VDMPNVITGN
 101 NKMFQHEWN KHGICSVSRF DQAAFYOMAI NMRNSIDLLS ALRVGGVVPN
 151 GRSKARQRV SAIRALQKE PVLRCGTGR QSRLEIYMC FDDGVTLLN
 201 CNPANSKPN SFTF

11AA_SEQUENCE 1.0
 ID Q04381 PRELIMINARY; PRT; 521 AA.
 AC Q04381;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE PROTON/PHOSPHATE CO-TRANSPORTER.
 GN ATP14.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-97375319; PubMed-9235602;
 RX Lu Y.-P., Zhen R.-G., Rea P.A.;
 RT "AtPT4: A fourth member of the Arabidopsis phosphate transporter gene
 family (Accession NO. U97546) (PGR97-082).";
 RL Plant Physiol. 114:747-747(1997).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: U97546; AAB69122.1;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.

KW Transmembrane.
 SO SEQUENCE 521 AA; 57356 MW; 62DF9254EC6942D5 CRC64;
 O04381 Length: 521 May 30, 2002 09:32 Type: P Check: 4865 ..

1 MADQQLGVK ALDVARTOLY HETAIVIAMG GFPTDAYDLF CVALVTKLIG
 51 RLTYFNPTSA KPSLPPHVA AAVNGVALCG TLAGOLFPGW LGDKLGRKKV
 101 YGTLIMMIL CSVASGLSIG NSAKGVMTLL CFFRWLGGF IGGDYPISAT
 151 IMSEYANKKT RCAPFAVFA MOGVRILAGG FVALAVSIF DKKEPSPTYE
 201 QDRLESTPPQ ADVIMTIYV FGLPALALTY YWRKMPETA RYALVAKNI
 251 KQATADSKV LQTDLEER VEDDYKDRK NYGLFSKEFL RRRGLHLGT
 301 TSTWFLDLIA FYSQNLFOKD IPSAIGWIPK AATNNAIHEV FKIRAOQLI
 351 ALCTVPGYW FTVAFIDIG RFAQLMGPF MMTVFMFAIA FYNHMLPD
 401 NRIGFVWYS LTFEFANFG NATTFIVPAE IFPARLSTC HGISATGKA
 451 GAIVGAFGL YAAQPODKTK TDAGYPPGIG VKNSLIMLV INEVGMLTFP
 501 LVPEPKGKSL EELSGAEVD K

11AA_SEQUENCE 1.0
 ID Q43650 PRELIMINARY; PRT; 540 AA.
 AC Q43650;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE INORGANIC PHOSPHATE TRANSPORTER 1.
 GN STPT1.
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-CV. DESIREE; TISSUE=ROOT;
 RX MEDLINE-97246321; PubMed-9090882;
 RA Leggewie G., Willmitzer L., Riesmeier J.W.;
 RT "Two cDNA's from potato are able to complement a phosphate uptake
 deficient yeast mutant: Identification of phosphate transporters from
 higher plants.";
 RL Plant Cell 9:381-392(1997).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: X98890; CAA67395.1;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.
 SO SEQUENCE 540 AA; 59108 MW; C7983E14933763BF CRC64;
 O43650 Length: 540 May 30, 2002 09:32 Type: P Check: 8926 ..

1 MANDQLVLA LDVAKTOLYH FTAIVIAMG FETDAYDLFC ISWTKILGR
 51 IYHHNDALK PGSLLPPNSA AVNGVAFCT LAGOLFPGWL GDMKGRKKV
 101 GMTLMINVC SINSGLSFH TPKSVMITLC FFRFWLGGI GGDYPLSATI
 151 MSEYANKRTR GAPFAVFA OGFGLAGM VAIIVSSAFK GAPPAPEV
 201 DALASTVSA DFWRIILM GAIPAGLFYV WRMKMPETAR YFALVAKNLK
 251 QAAADMSKVL QVEIEAPEK VAISVANCA NEGLFSKEF LRRHGLHLG
 301 TASTWFLDTI AFYSQNLFOK DIFSAIGWIP PAQTMNALEE YKIRAOQL

351 IALCSTVPGV WETVAFIDRI GRFAIQLMGF FEMTVEFAL ALPYHMTLK
 401 DNRIGFVVMY SLTFFPANFG PNATTFVVA EIPPARLRST CHGISAAAG
 451 AGAMVAGFGE LYAAPTDPK KTDAGYPAGI GVRNSLIVLG CVNFGMLFT
 501 FLVPSKSGKS LEEMSRNENG EETVAEMRA TSGRTVLFK
 !!AA SEQUENCE 1.0
 ID 041479 PRELIMINARY; PRT; 527 AA.
 AC 041479;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE INORGANIC PHOSPHATE TRANSPORTER 2.
 GN Stpt2.
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OC NCBI_TaxID=4113;
 OX NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. DESIREE;
 RX MEDLINE=97246321; PubMed=9090882;
 RA Leggewie G., Willmitzer L., Riesmeyer J.W.;
 RT "Two cDNA's from potato are able to complement a phosphate uptake
 RT deficient yeast mutant: Identification of phosphate transporters from
 RT higher plants.";
 RL Plant Cell 9:381-392(1997).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: X88891; CAA67396.1; -;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 DR Transmembrane.
 KW SEQUENCE 527 AA; 5796 MW; 2F6D8FD9FA50C85C CRC64;
 SO
 Q41479 Length: 527 May 30, 2002 09:32 Type: P Check: 2263 ..

1 MAVEDNNLOV LNALDLAKTO LYHFTAIITA GMGFETDAVD LFSISLVTKL
 51 LGRLYTKPD LKPGTLPPA RVGLRHGAL VGTLAGQLF GCARLAKMR
 101 KRYVGTILV MVMCSVASGL SLGNTPKVVM TTLCFRFVL GKGIGDPL
 151 SATIMSEVAN KKTGAFIAA VFAMQGGIIL FSGIVALITA AGPDHAYKAP
 201 TEAENAAVST VPQADYIURI ILMFGLPAA LTYWMKMP ETARYTALVA
 251 KQAKRAQOM GRVLOVEIES EEAKEIQISR DETNQGIFS WEFVRRHGLH
 301 LFGTGTWFL LDIAFYSONL FQKDVFSAVG WIPKAPLMA VOELXIARA
 351 QTLIALCSTV PGYMTVAFI DIIGRAIQL MGFEFMTVF FAIAIPYHMM
 401 TLEANNIGFI VMYSLTFFPA NEGNATTFV VPKEIPARL RSTGCGISAA
 451 AGKAGAIIVA YGELYAOSK DPMKTDAGY AGIGIKNSLI VLGFTNALGM
 501 VCTFCVPSK GKSLSEASOE TISTGEA
 !!AA SEQUENCE 1.0
 ID 082003 PRELIMINARY; PRT; 538 AA.
 AC 082003;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PUTATIVE PHOSPHATE TRANSPORTER.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. EARLY MECH; TISSUE=ROOT;
 RA Tahiri-Alaoui A., Avrova A., Antoniw J.F.;
 RT "A phosphate transporter from tomato roots up-regulated during
 RT arbuscular mycorrhiza colonization.";
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: Y16125; CAA76075.1; -;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 DR Transmembrane.
 KW SEQUENCE 538 AA; 58829 MW; F93PB158A4A87115 CRC64;
 SO
 082003 Length: 538 May 30, 2002 09:32 Type: P Check: 5788 ..

1 MANDLOVLNA LVYAKTQLYH FTALVIAGMG FFDVAIDLFC ISKVTYLLGR
 51 LYYHHDGALK PGSLPPNVSA AVNGVAFCGT LAQQLFPGWL GPKMGRKKY
 101 GMTIMIVIC SIAGLSFGH TPKGWMTLIC FFRFWLFGFI GGDYPLSATI
 151 MSEYANKKTR GAFIAAVFAM QGFGILAGM VAIIVSAFK GAEPAPAYEV
 201 DAIGSTVPPQ DEFWRIIMF GAIPAGLTY WRMKMPETAR YVALVAKNLK
 251 QAANDSKVL QVEIEAPEK VTALIKGKA NDFGLFTKEF LRRHGLHLG
 301 TASTWELDI AFYSQNLPRK DIFSAGWIP PAQTMALAE VKIARAQTL
 351 IALCSTVPGV WETVAFIDRI GRFAIQLMGF FEMTVEFAL ALPYHMTLK
 401 DNRIGFVVMY SLTFFPANFG PNATTFVVA EIPPARLRST CHGISAAAG
 451 AGAMVAGFGE LYAAPTDPK KTDAGYPAGI GVRNSLIVLG CVNFGMLFT
 501 FLVPSKSGKS LEDLSRENEG EETVAEMIRA TSGRTVLFV
 !!AA SEQUENCE 1.0
 ID 094DB8 PRELIMINARY; PRT; 555 AA.
 AC 094DB8;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PUTATIVE PHOSPHATE TRANSPORTER.
 GN P0694A04.13.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=4530;
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0694A04.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF003294; BAB63731.1; -;
 DR F293412D5F6A3B76 CRC64;
 SO SEQUENCE 555 AA; 60087 MW; F293412D5F6A3B76 CRC64;
 094DB8 Length: 555 May 30, 2002 09:32 Type: P Check: 3975 ..

1 MADADGSMNL AVLDALDSAR TQWTHMKAIV IAGMGFTDA YDLFCISTYS
 51 KLIGRLITYOP DGSTDSKPGA LSTANNKVI GVALVGTLMG QLVGFGGDK
 101 LGRKRYGYVT LILMAACAIG SGLSFGSSRK AVIGTLCFFR FWLGGIGGD
 151 YPLSATIME YSNKTRRAF IAAVFAMGV GIIFAGLVSM IYSSIFLTYN

201 KAPSYKGNHD LSHQMPAADY VVRIVLMIGA FPAALFYWR MKMBETARYT
 251 AIIDGNNAQA ANDMOKVLST EIEAEQEKLA KFNANNPYL LSMFARRHG
 301 LHLIGTTTW FLDDIATYSQ NLTQKDIFPA MGLISGAAY NALTEMQIS
 351 KASFLVALLG TPEGYWTVTA LIDKMGRYMI QLIGFFMMSM FMLAMGILYD
 401 YLKTNNHFLFG LLYALTFEFA NFGPNSTTV LPALFPTRV RSTCHASIAA
 451 AGKAGAIYVA FGIOKLTYSN QVKSIRKALI ILSTNNMGF FFFTLVPEYM
 501 GRSLEEISGE DGNLTAGGGG APAANAGV VSASDVSRDE KFPASSTEMQ
 551 TSMHA

11AA_SEQUENCE 1.0
 ID 0949N0 PRELIMINARY; PRT: 549 AA.
 AC 0949N0;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHEICAL 62.4 KDA PROTEIN.
 OS Arabidopsis thaliana (mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,
 RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L.,
 RA Caranci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koeseema E., Lam B.,
 RA Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shino P., Southwick A., Tracy S.E.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene F13M14.22/AT3g10500 (GI:12322791).";
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY051003; AAK93680.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 549 AA; 62356 MW; 89903EF6B7AFB5F5 CRC64;

0949N0 Length: 549 May 30, 2002 09:32 Type: P Check: 1025 ..

1 MGRSSTVSLA PGRFPHPTDE ELVRYLKRK ICNRPFKDA ISVNDYKSE
 51 PMDLPRSRL KSRDLWYFF SMLDKRYNG SKTNBATEMG YWKTGKDRD
 101 ILNCSKVGM KTLVYHKR APRGERTNMV MHEYRLVDD LDKGVHQA
 151 FVLGRITOKS GSGPKNGEY GAPVEEWE EEDMTFPPD QEDGSEBHV
 201 YVHMDDIDOK SENFVYDAI PIPLNFHGE SSNNVETNYS DSINYIOGTG
 251 NYMDSGYFE QPAEYKQD KPIIRDRGS LQNEGICGV QDKHSETLOS
 301 SDNFICDTS CYNDFPESN YLIGFAFLDP NSNLENDGL YLENDLSST
 351 QODGFDEEDY LTFEDETDP SOLMGNDYF FDQELFOEV ETKELEKEET
 401 SRSKHVEEK EKDEASCSQ VDADATEFEP DYKYPPLKKA SHMGAIAPAP
 451 LANSEFPTK DAIRLHAQ SSGSVHTVAG MTTIDSMMG WSYCKNENLD
 501 LILSLGVQG NTAPEKSGS SAMAMLEFMC FWILLVSVF KVSILVSR

11AA_SEQUENCE 1.0
 ID 0946Y1 PRELIMINARY; PRT: 520 AA.
 AC 0946Y1;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CERAMIDE GLUCOSYLTRANSFERASE.
 OS Gossypium arboreum.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Malvales; Malvaceae; Gossypium.
 OX NCBI_TaxID=29729;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21424002; PubMed=11443131;
 RA Leipeit M., Warnecke D., Zahring U., Ott C., Muller F., Hude B.,
 RA Heinz E.;
 RT "Glucosylceramide synthases, a Gene Family Responsible for the
 RT Biosynthesis of Glucosylolipids in Animals, Plants, and Fungi.";
 RL J. Biol. Chem. 276:33621-33629(2001).
 DR EMBL; AF367245; AAK73021.1; -;
 KW Transferase.
 SQ SEQUENCE 520 AA; 58616 MW; CDB77062770BE715 CRC64;

0946Y1 Length: 520 May 30, 2002 09:32 Type: P Check: 9530 ..

1 MSAALDPYDM LFLSLKAFR SPLAVVQIO GCVICITLAI GMATPAAYRN
 51 REINRMKDM KCGNSFAFLC HDINELEHTN QVNLBRVTV MPLGFGEHN
 101 LHMNSQITS LYGGPLEFLF VVESTEDPAY HAVSRILRDF KDVDAKITV
 151 AGLSTTCSOK IHNDLVEGR MHKDTKYVLF LDDVRLHNG SIGALTAME
 201 KNPEIFIQTG YPLDLPSSGL GSICYIYEHM PCSMGFATGG KTFPLMGCM
 251 MMQADDFRRD NYGVVSGLRD GGYSDMTILA AIAGAKRLI TSPVAVPEH
 301 PLASDLSRSR YWNLKRGTF VLESYSRVN WLMNGLPSF HCYLSMGVA
 351 PYMAAVHIA AALQIYKGY SYEETCTTS GLLASCLAI CTLTELLSMW
 401 NLTRIEVQLC NMLSPAPKL SLDYWNLSL FVALLVDFNL YPISAFRSHF
 451 SOSIMWSGIR YHLKNGKINK IERNKGRGR FFDLGKHLX GKKAAPPKAS
 501 FLSSIASRLC QMHQPKREV

11AA_SEQUENCE 1.0
 ID 0945E6 PRELIMINARY; PRT: 538 AA.
 AC 0945E6;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PHOSPHATE TRANSPORTER.
 GN OSPT1.
 OS Oryza sativa subsp. indica.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzae; Oryza.
 OX NCBI_TaxID=39946;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. IR36; TISSUE=ETIOLATED SHOOT;
 RA Kim D.-H., Nam J., Jeong S.-J.;
 RT "Oryza sativa OSPT1 mRNA for phosphate transporter.";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF416722; AAL11542.1; -;
 SQ SEQUENCE 538 AA; 58686 MW; 240FF49227F35A0 CRC64;

0945E6 Length: 538 May 30, 2002 09:32 Type: P Check: 7000 ..

1 MAGELVLNA LDSAKTQMYH FTAIVAGMG FFTADYDLS ISLVTKILGR
 51 IYFNPASKS PGSUPPNVSA AVNGVAFCGI LAGOLFPGWL GDKKGRKKYV
 101 GNTLMLNVIIC CLASGLSFGS SAKGVNATLC FFRWLGFGEI GGDYPLSATY

```

151 MSEYANKRTR GAFIAAVFAM QGFCNLTCGI VAIIVSAFK ARPDAPAYRD
201 DRAGSTVPOA DYAMRIVIML GAIPALLTYY WRMKMPETAR YTALVAKNAK
251 QAAADMTQVL NVEIVEDEEK ADEVARRBOF GLFSGROLRR HGRHLGTTY
301 CWRVLDIAFY SSNLROKDIY TAVOULPKAD TMSALEMFK ISAOQLVAL
351 CGTIPGYWFT VEFIDIGRF VIQLGGFFEM TAFMLGLAVP YHHMTPGTT
401 SGFVWYAF FEFANEGPNS TTFIVPAEIF PARLRSTCHG ISAAAGKAGA
451 IVSGFGLYA AOSTDASKTD AGYPPGICVR NSLEFLAGCN VICGFFTELV
501 PESGKSLEB LSGENEDDD VPEAPSTADH RTAPAPPA

11AA_SEQUENCE 1.0
ID 0944R4; PRELIMINARY; PRT; 519 AA.
AC 0944R4;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE AT2G19880/FeF22.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RA SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carlini P., Dale J.M., Goldsmith A.P., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Kallin Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,
RA Becker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF24585; AAL1579.1; -
SO SEQUENCE 519 AA; 58713 MW; 7257989790DF4BA5 CRC64;

0944R4 Length: 519 May 30, 2002 09:32 Type: P Check: 3251 ..

```

```

DE INORGANIC PHOSPHATE TRANSPORTER (FRAGMENT).
GN Pt.
OS Thineopyrum intermedium.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Thineopyrum.
OX NCBI_TaxID=85679;
RN [1]
RA SEQUENCE FROM N.A.
RA Davies T.G.E., Ying J., Xu Q., Li Z., Gordon-Weeks R.;
RT "Analysis of high-affinity phosphate transporter expression in alien
RT translocation lines of Chinese winter wheats.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ13964; CAC8705.1; -
FT NON_TER 1
FT TER 125
SO SEQUENCE 125 AA; 13348 MW; BCF6E0C47CB27122 CRC64;

0943Y7 Length: 125 May 30, 2002 09:32 Type: P Check: 3376 ..

1 DLFICIALVTR LGRIYYTDP ALNEGHLPA NVSAAVGVA LCGTLAQLF
51 FGWLGDKLGR KSYVGFTLLI MVLCIASGL SFGHEAKGM GLCFFRFML
101 GFGVGGYPL SATMSEYAN KTRG

11AA_SEQUENCE 1.0
ID 093WR0; PRELIMINARY; PRT; 525 AA.
AC 093WR0;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE INORGANIC PHOSPHATE TRANSPORTER.
GN Pt2.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OX NCBI_TaxID=4565;
RN [1]
RA SEQUENCE FROM N.A.
RA TISSUE-PHOSPHATE STARVED ROOTS;
RA Davies T.G., Ying J., Xu Q., Li Z., Gordon-Weeks R.;
RT "Analysis of high-affinity phosphate transporter expression in alien
RT translocation lines of Chinese winter wheats.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ344240; CAC6985.1; -
SO SEQUENCE 525 AA; 57447 MW; E7034BDAF73F7E3 CRC64;

093WR0 Length: 525 May 30, 2002 09:32 Type: P Check: 9360 ..

1 MAFVQLNLK ALDVAKTQLY HEKAVVIAGM GFETDAYDLF CIALVTKLLG
51 RIYYTDPALN EPGHLRPNVS AAVNGVALCG TLAGLFFGW LGGDKGRKSV
101 YGFTLLIMVL CSTIASGLSG HEAKGMGTL CFFRFNLGFG VGDDYDLSAT
151 IMSEYANKKT RGTETAAVFA MGGFGLFCT IVTIISSAF RHAFAAPPFY
201 IDAASISGE ADVWRIIVM FETIPALTY YMRKMBETA RYALLAGNT
251 KQATSDMSKY LNKELISENV OGERATGDTW GLPSRQMKR HGCHLLATTS
301 TWPLLDVAFY SONLFQKDF TKIGWIPPAK TMNALLELYR IARAQALIAL
351 CGTVPGYWFT VAFIDIGRF WQLMGFTMM TIFMLALAI PVDLVAPGHH
401 TGPVLYLGLT FEFANEGPNS TTFIVPAEIF PARLRSTCHG ISAAAGKAGA
451 IIGAAGFLYA SODQKRPFG YSRGIGMNA LEVLAGTNFL GLFSLILVBE
501 SKGKSLEELS KENVGDDDTI VPTGV

```

11AA_SEQUENCE 1.0
 ID Q93W09 PRELIMINARY; PRT; 525 AA.
 AC Q93W09;
 DT 01-DEC-2001 (TREMBLERel. 19, Created)
 DT 01-DEC-2001 (TREMBLERel. 19, Last sequence update)
 DE INORGANIC PHOSPHATE TRANSPORTER (FRAGMENT).
 GN
 OS *Triticum aestivum* (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC *Triticum*.
 OC NCBI_TaxID=4565;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE-PHOSPHATE STARVED ROOTS;
 RA Davies T.G., Ying J., Xu Q., Li Z., Gordon-Weeks R.;
 RT Analysis of high-affinity phosphate transporter expression in alien
 RL translocation lines of Chinese winter wheats.*;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ344241; CAC69856.1; -;
 FT NON_TER 1
 FT SEQUENCE 525 AA; 57481 MM; 367463B2P9D7B46A CRC64;
 SQ

Q93W09 Length: 525 May 30, 2002 09:32 Type: P Check: 9330 ..

1 MATQGLNVLK ALDYAKTQLY HFKAVIAGM GFETDAYDLF CIALVTKLIG
 51 RIYTDPLN EPGLPANVS AAVNGVALCG TLAGLFFGWL LGDLGRKSV
 101 YGFTLLMVL CSIASGLSFG HEAKGMGTL CFFRMLFGV VGGDYPISAT
 151 IMSEYANKKT RGTFLAIVFA MQGFILFCT IYTIYSSAF RHAPAPPFY
 201 IDAASISPE ADYWRITVM FGTIPALTY YWRKMPETA RYTLALIGNT
 251 KQATSDMSKV LNKISEENV OGERATGDTW GLFSROFLKR HGVLHATTS
 301 TWFLDVAFY SQNLFOKDIK TRIGWIPAK TNNALIELYR IARAQIAL
 351 CGYVPGWFT VAFIDIGRF WIQMGFTMM TTFMLALIP YDYLVKRGHH
 401 TGFVLYGLT FEFANFGPNS TTFIVPAEIF PARLRSTCHG ISAAATGRAGA
 451 IIGAFGLYA SQDQKPEKG YSRGIGMRNA LFLVAGTNFL GLFSLVPE
 501 SKGKSLLELS KENVGDDDT VPTGV
 11AA_SEQUENCE 1.0
 ID Q93W08 PRELIMINARY; PRT; 525 AA.
 AC Q93W08;
 DT 01-DEC-2001 (TREMBLERel. 19, Created)
 DT 01-DEC-2001 (TREMBLERel. 19, Last sequence update)
 DE INORGANIC PHOSPHATE TRANSPORTER (FRAGMENT).
 GN
 OS *Triticum aestivum* (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC *Triticum*.
 OC NCBI_TaxID=4565;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE-PHOSPHATE STARVED ROOTS;
 RA Davies T.G., Ying J., Xu Q., Li Z., Gordon-Weeks R.;
 RT Analysis of high-affinity phosphate transporter expression in alien
 RL translocation lines of Chinese winter wheats.*;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ344242; CAC69857.1; -;
 FT NON_TER 1
 FT SEQUENCE 525 AA; 57377 MM; 81ECBC50849F6689 CRC64;
 SQ

Q93W08 Length: 525 May 30, 2002 09:32 Type: P Check: 8781 ..
 1 MATQGLNVLK ALDYAKTQLY HFKAVIAGM GFETDAYDLF CIALVTKLIG
 51 RIYTDPLN EPGLPANVS AAVNGVALCG TLAGLFFGWL LGDLGRKSV
 101 YGFTLLMVL CSIASGLSFG HEAKGMGTL CFFRMLFGV VGGDYPISAT
 151 IMSEYANKKT RGTFLAIVFA MQGFILFCT IYTIYSSAF RHAPAPPFY
 201 IDAASISPE ADYWRITVM FGTIPALTY YWRKMPETA RYTLALIGNT
 251 KQATSDMSKV LNKISEENV OGERATGDTW GLFSROFLKR HGVLHATTS
 301 TWFLDVAFY SQNLFOKDIK TRIGWIPAK TNNALIELYR IARAQIAL
 351 CGYVPGWFT VAFIDIGRF WIQMGFTMM TTFMLALIP YDYLVKRGHH
 401 TGFVLYGLT FEFANFGPNS TTFIVPAEIF PARLRSTCHG ISAAATGRAGA
 451 IIGAFGLYA SQDQKPEKG YSRGIGMRNA LFLVAGTNFL GLFSLVPE
 501 SKGKSLLELS KENVGDDDT APTGV
 11AA_SEQUENCE 1.0
 ID Q9D6E1 PRELIMINARY; PRT; 179 AA.
 AC Q9D6E1;
 DT 01-JUN-2001 (TREMBLERel. 17, Created)
 DT 01-JUN-2001 (TREMBLERel. 17, Last sequence update)
 DE 2900075G08RIK PROTEIN.
 GN 2900075G08RIK.
 OS *Mus musculus* (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=HIPPOCAMPUS;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamane A. I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombertus P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.*";
 RL Nature 409:685-690(2001).
 DR EMBL; AK013789; BAB28996.1; -;
 DR MGD; MGI:1920228; 2900075G08RIK.
 SQ SEQUENCE 179 AA; 21105 MM; B738709483743E82 CRC64;
 Q9D6E1 Length: 179 May 30, 2002 09:32 Type: P Check: 164 ..

1 MATFAACY MLSLVLCAL IFFAIWHIA FDELRTDKS PIDOCNPYA
 51 RERLRNIRI CELLRKRGV PPGRRKRR ERGOOLVPE YSISLPCIM
 101 FLCAQEWLTL GLNPLFLYH FWRYHCPAD SSELAVDPV VMNDTLISC
 151 OKANCKLAF YLSFFYLY CMITLVSS

```

!!AA_SEQUENCE 1.0
ID 09CZ13 PRELIMINARY; PRT; 139 AA.
AC 09CZ13;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 2310040A02RIK PROTEIN.
GN RNC OR 2310040A02RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whitlaker C., Wilmink L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsuki S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
RL EMBL: AK012585; BAB2835.1; -.
DR MGD: MGI:1913275; Rnac
SQ SEQUENCE 139 AA; 15596 MW; 6FBFBAC137FAC0A CRC64;

09CZ13 Length: 139 May 30, 2002 09:32 Type: P Check: 6690 ..

1 MATQANSLSY AGCNFLRHWL VLSLISGRPV KIRVRARD NPGLRDFEAS
51 FTLLKRTIN GSRIETNQTG TLLYQPELL YGGSVEHDCS VLKIGLYLE
101 ALLCLAPFMK HPLKIVLRGV TNDQVDPESVD VLKATRSLEY

!!AA_SEQUENCE 1.0
ID 063363 PRELIMINARY; PRT; 475 AA.
AC 063363;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE RAT NOEYTHROID ALPHA-SPECTRIN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88094373; PubMed=3336352;
RA Leto T.L., Fortugno-Erlson D., Barton D.E., Yang-Feng T.L.,
RA Francke U., Harris A.S., Morrow J.S., Marchesi V.T., Benz E.J.Jr.,
RT "Comparison of nonerythroid alpha-spectrin genes reveals strict
RT homology among diverse species.";
RL MOL. Cell. Biol. 8:1-9(1988).
DR EMBL: M19726; AAA1678.1; -.
DR HSSP: P07751; IAUJ.
DR InterPro: IPR002017; Spectrin.
DR Pfam: PF00435; Spectrin; 5.
DR SMART: SM00150; Spect; 4.

```

```

FT NON_TER 1 1
SQ SEQUENCE 475 AA; 54713 MW; 3E0468C6FB6E03B7 CRC64;

063363 Length: 475 May 30, 2002 09:32 Type: P Check: 7377 ..

1 RDMDEESWI KEKKLLVSGE DYGRDITGVQ NLRKKHRL EGRLEPQAI
51 QGVLDGSKRL IDNTIGREE IOORPEQVE HMKELKQIAA ARGORLEESL
101 EYQGFVANVE EEFAMINEKM TLVASEDYGD TLAIIQGLK KHEAFEDFT
151 VHKRVADV C TNQODLIKNN NHHEENISK MKGLNGKYSV LKRAAQRKA
201 KLDENSAFLQ FNRKADVES WIGERENSLK TDYGRDLS VQTLTKQET
251 FDAGLQAFQO EGIANITALK DQLAAKHQ SKATEARHNS LMRKWQLFA
301 NSATRRKKLL EAQSHRRKVE DLFLAPAKKA SAPNSWEENA EEDLTDVRC
351 NSLEIRKALR EAHDAFRSSL SSAQDFNQL AELDRQIKSF RVASNPYTF
401 TMEALEETWR NLQKIKERE LELQKEGRRO EENKLRQEF AQHANAFHW
451 IQERTTILD GSCIGRRVCN SGISA

!!AA_SEQUENCE 1.0
ID 09DCJ5 PRELIMINARY; PRT; 172 AA.
AC 09DCJ5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 0610033L03RIK PROTEIN (RIKEN CDNA 0610033L03 GENE).
GN 0610033L03RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whitlaker C., Wilmink L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsuki S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX Strausberg R.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK002744; BAB2232.1; -.
DR MGD: BC012416; AAH12416.1; -.
DR MGD: MGI:1915625; 0610033L03RIK.
SQ SEQUENCE 172 AA; 19992 MW; 7DF62AB86B5BF684 CRC64;

09DCJ5 Length: 172 May 30, 2002 09:32 Type: P Check: 5548 ..

1 MPTIVELPTL ELKVEEVAV SSAVILKAAAH HYGACDCKTN KEPMICRWEE

```

51 KDPKRCLEKES KLVNGCALNE FROIKSHCAE PFTEYWTCLD YSNMQLFRHC
 101 RQOQAKFDQC VIDKLGWVR DUGOLSKYTK VKTRPLPEN PYHSRAPER
 151 NPVIEGDLKP AKHGTRFFEM TV

11AA_SEQUENCE 1.0
 ID 091YD4 PRELIMINARY; PRT; 1507 AA.
 AC 091YD4;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE TRANSPARENT RECEPTOR POTENTIAL CHANNEL 7 (FRAGMENT).
 TRP7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RA Keshuba V.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ344343; CAC69081.1; -
 KW Receptor.
 FT NON_TER 1507
 FT SEQUENCE 1507 AA; 172364 MW; 86E552DE25939785 CRC64;
 SQ

091YD4 Length: 1507 May 30, 2002 09:32 Type: P Check: 1512 ..

1 MESIDRRRTG SEQEGFGVO SRRATDLGMY PNLRRSNSL CKSRRLFCSE
 51 SSEQENLSS WIPENIKKKE CYFVESSKL SDACKVOC GYHGHLEV
 101 AIKPHITQOK EMDPKKHVOE MPTDAPGDIY FTDLSQKYGK YRVSDTPS
 151 SVIYQMTQW WGLDVPNLLI SVTGAKNFN MKLRKLSIFR RGLVKAQTT
 201 GAWIITGSH TGVMQVGEA VPDFSLSSC KEGEVITIGV ATMGTHNRE
 251 GLHPMGGFP AEYMDDEGO GNLTCDSNH SHFLVDDGT HGQGYEIPL
 301 RTKLEKISE QTKERGVAI KPIVCVYLE GGPGLHTIY NAINNGTPCV
 351 IVESGSHVAD VIAQVATLPY SEITISLIQO KLSIFQOMF EFTTENIVE
 401 WTKIQUOIVR RROLTITFRE GKDGOQDDVY AILQALKAS RSODHGHEN
 451 WDHOQLKAVA WNRVDIARSE IFTDEMOKRP ADLHPMTAA LISNKEPFR
 501 LELENGVRLK EFTVMDTLIC LYENLEPSC L FHSKLQKVLA EEOQLAVASA
 551 TPRLMHMIVA QVIRELIGDS TOLLTPRPY TDRRLSTV, PHIKNOGV
 601 SLRSLYRST GHVFTTIDPY RDLLIMAYIO NHRELAGIIV AOSODCTAA
 651 LACSKILKEL SKREEDTDS EEMALADEF EHRAIGVTE CYRKDEERAQ
 701 KLVYVSEAW GKTTCLQAL EAKDKFVSH GGIOAFITKV WNGOLCYDNG
 751 LMRILICMLA FPLLTGFS FREKRLQALC RPARVAFEN APYVIFHMI
 801 LSYFALICLF AYVLWDFQO SPSWCEYLY LMLSLVCEE TROYLFDPD
 851 GCGLMKASL YFSDFWNKLD VCAILLFIYV LTRCLIPATL YPGRIILSD
 901 FIMECLRLMH IFTISKTLG KLIIVKRMK DVEFFELLA VWYVSEGVAK
 951 QALIHNESR VDWIFRGVY HSYLTIFGOI PTYIDGVNES MDCCSPGTD
 1001 PYKRCESD WTGOAPAFPE WLTVTLLCLY LLFANILLN LLAMENYTF

1051 QEVOEHTDOI WKFORHDLIE EYHGRPPAP PLILSLHQL LIRIYLVKIP
 1101 AKRHQOLKNK LEKNEETALL SMELYLKENV LONQOYQOK RPEKIDIS
 1151 EKVDTMDDL DMOVKRSGS TEORLASLEE QYOVTRALH WYTTLKDSG
 1201 FSGGAGALT APQRADEPD AELSIRKVE EPGDGHVSA RHLLYPNARI
 1251 MRFPVNEKV PMAAEFLIYD PEYTAEKDV ALTDPVGDTA EPLSKISYN
 1301 VDGPTDRSF HGVYVVEYGF PLNPMGRICL RGRSLSWFG PNHTLOPVVT
 1351 RWKRNOGAI CRKSVRKMLE VLVMKLPSE HWALPGSGRE PGEMLPKRLK
 1401 RVLRQEFVVA FETLLMGTE YKGYVDPR NTDNAWITV AVSIHPDON
 1451 DMELKRLLEN LHTHDKELT RDLKUSTEQ VVDRIPIYA NHTIILQVA
 1501 SLFGANF

11AA_SEQUENCE 1.0
 ID 056074 PRELIMINARY; PRT; 3005 AA.
 AC 056074;
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE GENOME POLYPROTEIN.
 OS Hepatitis GB virus A.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage: Flaviviridae.
 OX NCBI_TaxID=39112;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97126113; PubMed=8971037;
 RA Leary T.P., Desai S.M., Yamaguchi J., Chalmers M.L., Schlauder G.G.,
 RA Dawson G.J., Mushawar I.K.;
 RT "Species-specific variants of GB virus A in captive monkeys.";
 RL J. Virol. 70:9028-9030(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97437486; PubMed=9292019;
 RA Leary T.P., Desai S.M., Erker J.C., Mushawar I.K.;
 RT "The sequence and genomic organization of a GB virus A variant
 RT isolated from captive tamarins.";
 RL J. Gen. Virol. 78:2307-2313(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98120818; PubMed=9460920;
 RA Erker J.C., Desai S.M., Leary T.P., Chalmers M.L., Montes C.C.,
 RA Mushawar I.K.;
 RT "Genomic analysis of two GB virus A variants isolated from captive
 RT monkeys.";
 RL J. Gen. Virol. 79:41-45(1998).
 RN [4]

RP SEQUENCE FROM N.A.
 RA Erker J.C., Desai S.M., Leary T.P., Chalmers M.L., Montes C.C.,
 RA Mushawar I.K.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF023425; AAC40502.1; -
 DR HSP; P27958; 1HEI.
 DR MEROPS; S29.002; -;
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR002518; HCV_NS2.
 DR InterPro: IPR004109; HCV_NS3.
 DR InterPro: IPR000745; HCV_NS4a.
 DR InterPro: IPR001490; HCV_NS4b.
 DR InterPro: IPR002868; HCV_NS3a.
 DR InterPro: IPR002166; HCV_RdRP.
 DR InterPro: IPR000566; Lipoclin_cytfAB.
 DR InterPro: IPR001005; Myb_DNA_bind.
 DR Pfam: PF01538; HCV_NS2; 1.
 DR Pfam: PF02907; HCV_NS3; 1.
 DR Pfam: PF01006; HCV_NS4a; 1.
 DR Pfam: PF01001; HCV_NS4b; 1.

DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00998; HCV_RdRp; 1.
DR PROSITE: PS00213; LIPOCALIN; UNKNOWN_1.
DR PROSITE: PS00037; MYB_1; UNKNOWN_1.
KW Nonstructural protein: Polyprotein: RNA-directed RNA polymerase.
SQ SEQUENCE 3005 AA; 323793 MW; BB407D1ADFA8162 C6C64;

OS06074 Length: 3005 May 30, 2002 09:32 Type: P Check: 8233 ..

1 MEVLVLLK TALAGFIOP ASHCHAGY YLLNCCSDS EITAFEDGG
51 LVSGVGVCD GRCMDLYRPG VATRGHGG ELFGALMSTS SPRAVAAYVA
101 GTIGLGEPPS AAALLAALAR VEMTPRPVNT CLMDCDIALQ SEAMDLVEDM
151 RDTFWYKVV MTLPMRLAVG LVGMSFALM VVLLLEOR VVAVLILAM
201 AGSEVEPTGG CRCHVNGSIV AIPGAKPRDR PNATQVCVCP FQGMVLPAL
251 CAGLAMRNGD VSGTTRDLPV RCPESMHGRG SVMCWGSAH MSWRLATEVR
301 LMEQLGSGAL CHEFAVGTID RPHPATDVLG THGIPCASCV VDRRASMCGN
351 CTRDCMEKGT NKRLSFEAGG LGSRLTAEIM AHVVDGSTES KIATVDGERP
401 RYPSHSGVGT HHTAVVAARN YTVSDIGIYW HAIACPNPPS PSALPRLIPG
451 RPNACLTIC KKKGLNTAM QAPGKXFPDI ETECNMPRRS GVRCEGYAF
501 DFGSKTGTI RSRAGFOQIA RAGSMHPDRM LMTDYLFIIL VMKLAEARL
551 VPLVAVALYM WENNAEAT LRLHPVNI TTAAPMTWPA LPPYVVCPT
601 KAGAVGDALY MAACLAGNAT LGVYSSVWG AYGAAEAGAR GLMRWGGFR
651 SFCAGLAWLT NVGAYLPVE AAVGPEVSA PLVWAMEDS IAVFIIMAN
701 IMAYMSHIG VKLGLVAH LAKGALPYL LVAAXYTRHR HSYLGEVCY
751 SLDADSPSD WSMWALAAYV SWCLITLGLT THGRLKLE FSYWCRFHO
801 AVEMRLCSK VGRGRDMRV TAVWVAAGII FPREVNCSA ALTALAALD
851 SIDYLETLLI LTAQPARAA RLDSLTFLG DADLTTRAFV RLRECVTLF
901 QHGGQVSXGA AAILADLGYA LEPVSVTARD CIYVDAART LACGQREGL
951 PVVARRGDEV LVGVFPVSVA LPPGFVPTAP VVWQRLGF FSVYKTSMLG
1001 RDRREHGST VVLGTTSTRS MGTGVNGVMY TTFHGSNART LAGPVGPVC
1051 RMWSPSDOVA VYPLDSGASC LEPCKGTQS WMCIRNDGAL CHGRLSKVE
1101 LDLPTEISDF RGSSGSPILC DEGHVGMV SVLHRGVKVT GVRVYKPMET
1151 LPKDSQYKSE APPVPGKTGF TEAPLYLPTG SKGSTRIPLF YRKAGIKVLY
1201 LNPSTATVRA MGPYMEKISG QHPSIYCGHD TTAVSRTGS PLTYCTYGR
1251 MANPRYRLAG ADIVICDECH VTDPSTVLGM GRARLLAREC GVRLLFATA
1301 TTPGAPLAQH ESIKVEPLGV DGEVAFYGHK LPVERVYTRG HLLFGSKVE
1351 CNRLHAALST AGCNVAVYVR GNEOEIPAGD VCVCATDALS TGYTGFSFY
1401 TDCGLAVEEV VEWTLDPIT ISVRTTPAPA ELRAQRGRG GRGSGTYYY
1451 AMRASAPAGT LRSGLPMAAV EAGVAYMLE PDMTADILRA YDACPYTAI
1501 TASVGEALNF FSGLVPMRY POVAMAKSHQ HNPPLIVGQ RTWCOEANTA
1551 GPDDGPEMAG LAGTGPIPL CRMGARPPS VAPHHWDDL QARLGADEGY

1601 SPCTAGPILL VGIALAGAV LAHWTGSLV VTSNRVNGG NPLIQDSTRG
1651 VSTSAFYQLT VCVEGEQTPA DKKCAEAAYQ LXESTCGMGP MAASFCAAM
1701 KGYLDSMRTT AAAAVEKSDS LMRSCANNY CPTGCGATSA SAFFASLDIK
1751 PQAQDAIFT NGRSLVGLV AGYARRNPP LGVAAPFLMG MSAGHOVHR
1801 LAAALLGVG GPMIGTPSYG LAMGAYFAG GSITSSMLSA IYAVLGMWEG
1851 AXNAASLTFD FLTGRELKD MWFVSCXAS PGASVAGVAL GLLMSMKKG
1901 VGEDWNRIL TLIPRGSVLP DGEVKSSEPT ERVSTILKRM SLGRWMTLV
1951 ERRELDETP CSSMLMDLID WLVRGRYIG RLKGWPSV RYLVGCTPG
2001 WGSWGEHG IEARCGGCI ITADVEBKL VDVHYSSRLC SNLKTIVY
2051 SAAGSDAEP EYPAGPALYQ IGVAEMVOLV RKDKTLVGA SSVYHLHIDE
2101 LRRAIRGPPM FVGGVSVME APLOOPLYV RAGQSVAFDD VRYSLPHTLA
2151 LPPPRPPPP LAMPPPPPP EQWTEEEEA DLRRARARAI EAVNERLPVP
2201 NPEAQAALD ALBEAAVSIL PHVGAIIQDD CSCESEFGGH FIPEPDAVDY
2251 PIGMEVQVG PLRDQARDLG DRLAVIAGRL ESILAEHPEA SLNTERRTWG
2301 ELIDITLADV ARLDVTCRSD TSGSSFQDIS LSDSEPTIV EGGLEKLEVR
2351 POPRFEDLI RPEGAKRLV TVKOSCCADR SATRAFLSL PIAAVTATLS
2401 EDLDTHTVSD STGRVLDPLE LLQNAVGDLI VACRRSSSV SYSYISGAP
2451 LNTGRHQPAR MTRPIGTHIT ADTKVYVTD PNQGERAAK VTIMGRSVY
2501 DAHYRGVSE VLDQAKTIKS PGWYDEAIA KYHRAAAGF GSNVYVGNMT
2551 TPAARAEVNN MLAKIKTQK VPTLVTRKE VFEQKTRKP PREICEPPID
2601 FRIAEKMILG DGLVAKGIL GKSYLEQYTP NQRYKLAVDL WRQKKHRAI
2651 TVDATCPDSS IDERDAVET EVFAASPNP DLYALDSY AEEPMSPPG
2701 VPLGVKRCRS SGLVLTSSAN SITCIKYKA AARAVGLVDP DFLIAGDDCV
2751 IYEDDEEDH ADALRYALGN YGYDCKPKTH ASLDTAESCS SYLAECNVGT
2801 ERYWWMLSTDH RKPILAAASE YSDPYSSALG TILMYPMHI VRVYLLPHIL
2851 IMAFRGGGTP DDLVCEVOG NHYSFPLRYL PEVLVSLHGP RCLRYTADST
2901 KTKKAGGAL RDLGMHTLAF YRKRAVNT RTLRGKGWG RLARALLMP
2951 GKKEHPPSIR STPGFKATP YEHHTWYS GKKPMYWDV KCLFGLYCAF
3001 LTALL

!!AA_SEQUENCE 1.0
ID 09DW56 PRELIMINARY; PRT; 655 AA.
AC 09DW56;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE PR139.
GN R139.
OS Rat cytomegalovirus (strain Mastricht).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_Taxid=79700;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MASTRICHT;

RX MEDLINE-20366325; PubMed-10906222;
 RA Vink C., Beuken E., Brugeman C.A.;
 RT "Complete DNA sequence of the rat cytomegalovirus genome.";
 RL J. Virol. 74:7656-7665(2000).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MASTRICHT;
 RA MEDLINE-20473137; PubMed-11018281;
 RA Gruijthuisen Y.K., Beuken E., Brugeman C.A., Vink C.;
 RT "Rat cytomegalovirus R89 is a highly conserved gene which expresses a
 RT spliced transcript";
 RL Virology 209:400-408(1995).
 DR EMBL; AF232689; AAF9234.1; -;
 DR InterPro: IPR003360; US22.
 DR Pfam: PF02393; US22; 1.
 SQ SEQUENCE 655 AA; 73553 MW; 652CEBF38F0664B47 CRC64;
 09DN56 Length: 655 May 30, 2002 09:32 Type: P Check: 4069 ..
 1 MMSLRGALRR DVERFLGREHH EMPTEWQDP TALEPTEDE DFGCVDDLD
 51 EKHIPLSRNP VYPPDDSLIK GVLGLSMSE GERGTRSGP LTGHNNMAL
 101 NPPDEIQT VSVRESREAS GHRHGGSDS DLGSDSGSD NEIIVGHS
 151 WKSTALTVSK EIVEEHEKN LSRREFATVL RDEFTLYVAL SNPEIPEPI
 201 RQSGRALTL RPEHMYLPL QPMNEIRVLE GVDLRAVIOC DVAARDDEL
 251 IPLGACAAARS PDLHYRETTPC AILMDTEGR FLYDAETEGFL FLAENLVEL
 301 AHKGLSACEP AYRDGATIS LPRRTAVRK ILSACVIGLE SYNAVTSVR
 351 GTAVILCDRA SGRERLIQIF DASELRKRP FSELDESGC AMMDYVERL
 401 AEBNVLGCI GTYDNGSGFV FMYDIYVLLG VTGAVYGFV EENDYRLAD
 451 DLSTLLRRCI SPEPNRPDR ALGELRLER ALCPHEREGT RGDEPVYRQ
 501 VARRDLRGL RMQLRTGPD DRCVPAHLE EAKROFRHPV SLPVCVVTG
 551 ESYQPKNSE QGYGVLAEG RNFRCPTRP DPTDEERRQ RELYQRMCP
 601 GPFVNOPEE TYTGVIARA ARLALQKSG TSLRLPAISR AATVSRGS
 651 PMSDR
 11AA_SEQUENCE 1.0
 ID 083896 PRELIMINARY; PRT; 595 AA.
 AC 083896;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE TERMINAL PROTEIN.
 OS Ovine adenovirus type 7.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Adenovirus.
 OX NCBI_TaxID=114430;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OAV287;
 RA MEDLINE-95297141; PubMed-7778275;
 RA Vreli S., Boyle D., Koehrhans R., Both G.W.;
 RT "Sequence of ovine adenovirus homologs for 100k hexon assembly, 33k,
 RT pVIII, and fiber genes: early region E3 is not in the expected
 RT location";
 RL Virology 209:400-408(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OAV287;
 RA MEDLINE-97080497; PubMed-8921842;
 RA Vreli S., Brookes D.E., Boyle D.B., Both G.W.;
 RT "Nucleotide sequence of ovine adenovirus tripartite leader sequence
 RT and homologues of the Iva2, DNA polymerase and terminal proteins.";

RL Gene 177:35-41(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OAV287;
 RA MEDLINE-97271311; PubMed-9126262;
 RA Xu Z.Z., Hyatt A., Boyle D.B., Both G.W.;
 RT "Construction of ovine adenovirus recombinants by gene insertion or
 RT deletion of related terminal region sequences.";
 RL Virology 230:62-71(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OAV287;
 RA MEDLINE-98277079; PubMed-9614874;
 RA Khatri A., Both G.W.;
 RT "Identification of transcripts and promoter regions of ovine
 RT adenovirus OAV287";
 RL Virology 245:128-141(1998).
 DR EMBL; U40839; AAD45952.1; -;
 DR InterPro: IPR003391; Adeno_terminal.
 DR Pfam: PF02459; Adeno_terminal; 1.
 SQ SEQUENCE 595 AA; 69781 MW; 026ADP6563066206 CRC64;
 083896 Length: 595 May 30, 2002 09:32 Type: P Check: 5502 ..
 1 MTAROIHOWQ QLTGOSQHTL RYLRLTLDL NRSIARSRTI HERGIKMANR
 51 FENYPTQLL DLRRGPYTS QPPEGEPP NLLIGFYVA RVLNQLFQD
 101 RYTSNISTKL YLSPISFERR MTWQILTDCS YSLNIGSYAR AIQNVENSQ
 151 TISQIOMAVL MDRLIGSLQI ADMQGFSAI SAONNRSEQ QHYQFSGAF
 201 MYLNQERPD YLIKLIKIK KALCKFLILS NMHNSCIDL LPFSEWIEL
 251 FIDFSQLEI PEARNRINK DLAIVTLGK GCMHGALTL RSGTVGLPF
 301 RLPRNRYRA ITETMRSGV DVIRRFIDL PVNRQRVEN VEETSSSPLT
 351 VEIEDEGFS DREPSDDL S REEENDEVIA SIYDLIONLE ELTPPEARSS
 401 NFNNGSEFF QLLIRFYEN RLTDPIQKW LYFFLEHIV ASTLYYLNL
 451 LVONRIARN IGIOFVQIIL RGNENGEDI FTRWNNRET QAFRLYSRI
 501 TRDFGIIEA NERSYFSTP EERDQLQDI DVEDSGSIE EYINOVNTDF
 551 SLDSEVEIAF RIKLSGIVY STNEVLRSF ERVREALNR WLQRO
 11AA_SEQUENCE 1.0
 ID 091GE6 PRELIMINARY; PRT; 243 AA.
 AC 091GE6;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHEICAL 26.9 KDA PROTEIN.
 OS Epiphyas postvittana nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OX NCBI_TaxID=70600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hyatt O., Dellow R.A., Olsen M., Caradoc-Davies K.M.B., Drake K.,
 RA Ward V.K.;
 RT "The complete sequence of the Epiphyas postvittana
 RT nucleopolyhedrovirus genome";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY043265; AAK85672.1; -;
 KM Hypothetical protein.
 SQ SEQUENCE 243 AA; 26869 MW; B5203A27C88CAEA0 CRC64;
 091GE6 Length: 243 May 30, 2002 09:32 Type: P Check: 8973 ..
 1 MGLNFOSY TRLPAPQIS LAVMSYVNT LCAYGAIVA YLSTATSFE

```

51 LQLEWVWL SLINGLINV TLEFOSKTE AHEIVEIKM LHAMFSSAL
101 VNLAIIDTAQ SASAVLVNN LIHCIALFL FVELTVLGH TLGTANYRY
151 TKACVLYVLL VTAAYTIIIL TAENKSSPL CNDLMAFSL TAFPMYIAYV
201 WAVEKAAGS VLORVOLISL YDPPSPSTNV KMEDMLANKQ MEI

11AA_SEQUENCE 1.0
ID Q9W7A2 PRELIMINARY; PRT; 380 AA.
AC Q9W7A2;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE FAST TROPONIN T ISOFORM.
GN TMTF.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OC NCBI_TaxID=93934;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=99292729; PubMed=10364205;
RX Bucher E.A., Dhooet G.K., Emerson M.M., Ober M., Emerson C.P. Jr.;
RT Structure and evolution of the alternatively spliced fast tropoin T
RT isoform gene."
RT J. Biol. Chem. 274:17661-17670(1999).
DR EMBL; AF13128; AAD4279.1; -.
DR InterPro: IPR001978; Tropoinin.
DR pfam: PR00992; Tropoinin; 1.
SQ SEQUENCE 380 AA; 43814 MW; 000A930E3D3689DC CRC64;

09W7A2 Length: 380 May 30, 2002 09:32 Type: P Check: 3716 ..

```

```

090ZD4 Length: 491 May 30, 2002 09:32 Type: P Check: 3654 ..
1 MAALDSGDA QKURFPMWL NAVSGLYRG LCIWDIPDRI FRIPWKHNR
51 KDVITSSDEVI FKAMAKASGR YEGNAEDPAK WKTFRCALR STHMFILLED
101 RSKCNDPPHK VVAASGVNP DRGSGPVAG ALQOQPOLLL NHHDLALENT
151 PTDSTEGVAA AALTQYDLDL LQSVLQHNCI SALGQPTLM AHNGDALPED
201 ALLPGQDGC LRGPOFQDMR QLEEPULIGN QPLTGGCGQ DGAGALPVS
251 ECALPAPSPA EELLFOSANP APPPAGDIG GLPPLDITI YRKGWYQOE
301 QVDDSRCVLA YQPLDPAVAE QRLVLEPSPA SLDPDQRRY TENULEVAGL
351 RLEBQAQQL ATRLKKCKYF WALSOQLBSG EPPNLNLRD QETTFIDERV
401 FCTELRDFRD SRRERSPDFT IFLCFGQCS STKPKESKLI LVKLVPQFCE
451 YWYEQVORG ASSLNSGNVS LQSDSFNLF ELIEQYHMQT D

11AA_SEQUENCE 1.0
ID 025915 PRELIMINARY; PRT; 417 AA.
AC 025915;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE COMPLEMENT LOCUS E (C3E3).
GN HP1361.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OC NCBI_TaxID=210;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khatkhat H.G., Glodek A.,
RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RT Nature 388:539-547(1997).
RL EMBL; AE000636; MMD08401.1; -.
DR TIGR; HP1361; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 417 AA; 47984 MW; B9D557BE7F7ED828 CRC64;

025915 Length: 417 May 30, 2002 09:32 Type: P Check: 8156 ..
1 MCGVFLSILL AINLYLEYLN YOKLDFSKPT SLAQIILQY PKTKDQKTYE
51 VKLQSKNMI FYTTIKKPKL NIQYHNAQF GKIPCSFLE SLKSCFPQY
101 SFSITRRQDF KSHWRHFIDS AHENALVGNL YRALFIDDSL NKDLRDRANA
151 LGINHLLAIS GFHLGISVS VFILSLFYT PLQKRYFPYR NAFYDIGLV
201 WVFPLGIVLL LDFLPSFFRA FLMLGLGFLA CFQSVRLISF KLIIACCTA
251 IALLPKILFS VGFLLSVCGV WYIFLELKH QIFPKTSSFL MRSFOAISIS
301 ALVEFLANLLI VHAFFPMFSP YQLFSIPLGL IFIYFFPLSL FLHAYVLGSL
351 LDRLLSPLT IPTISVSPDL WILGVHLFLT ILSARFFKYV LSMNVLSAGP

```

401 FLXCXYOYII MSPSLVG

11AA_SEQUENCE 1.0
ID P72685 PRELIMINARY; PRT; 290 AA.

DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE HYPOTHEORETICAL 32.1 KDA PROTEIN.

GN SL0696.

OC Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
ON NCBI_Taxid=1148;
ON [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugitani M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Mureki A., Nakazaki N., Naruo K., Okumura S.,
RA Shlomo S., Takeuchi C., Mada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
EMBL: D90899; BAA16692.1; -
DR InterPro: IPR003675; Abi. 1.
DR Pfam: PF02517; Abi. 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 290 AA; 32091 MW; AD481D2FA2AB4DD CRC64;

P72685 Length: 290 May 30, 2002 09:32 Type: P Check: 1310 ..

1 MPKSTVNFNF NQQAIAQLTA PMRIGVFIYM LAAAMLPFLV PLSAIAIDAN

51 LRSIVNGIL FLIFVLILF WSHWCYQPL SLKAYGYGL GWNRRQVEL

101 LRGLIGFSF TFGLEFIQGL LGMAYLAPAG DRLMTIING SLTGLGVALA

151 EELEFRGWL KELEGYGNK TSLASNAITF AVLHKLPLG EVITTLQOPF

201 ALVLLGISLG ITRRRHGDRL GHSIGLHGM VWAYIIVNG QLVYTERKP

251 AMYTGIDRNP LSGVMGIAGL CLLMLVING EKPMLKRLGF

11AA_SEQUENCE 1.0
ID Q9PAP6 PRELIMINARY; PRT; 544 AA.

AC Q9PAP6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHEORETICAL PROTEIN XF2449.

GN XF2449.

OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
ON NCBI_Taxid=2371;
ON [1]

RP SEQUENCE FROM N.A.
RC STRAIN=9AS5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Sampson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Agencio M.,
RA Alvares R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.R., Bonaccorsi E.D., Bordin S., Bove J.M., Blomes M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferrao J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.U., Honeisel J.D., Junqueira M.L., Kemper E.U., Kitajima J.P.,
RA Krieger J.E., Kurnae E.E., Laigret F., Lambais M.R., Leite L.C.C.,

RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Zeldman J., Zerbini J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
EMBL: AE004053; AAF85248.1; -
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 544 AA; 62881 MW; 45D840B178BBA046 CRC64;

Q9PAP6 Length: 544 May 30, 2002 09:32 Type: P Check: 144 ..

1 MICRRYVLVD SMNLFSSHW FVELLSVAL LLQCKQDSP SAKRQIIMVE

51 AKLPRPLPK DLPLPLDIE RRTFFQFMDT TNEVNGLRPD RYSPRPFASI

101 ASVGALNAV PIGIENGWIS RNOAIDRTLT TLKFLRDAM GPQRGCGAY

151 KGFYHFLDM QHGRYDSWV ELSSVDTALL MMGVLFVESY YDGEDVREKE

201 IROAIDKLXR RVDMVLYOOR KPLISMOWP ERGFIEDHM GYNAMAVYL

251 LALGSPTRPL EYVSWDEWR TYNKMGVFO GGEYLAFLG FGHQYTHWV

301 DERDIDQYM RENGIDYFLN SRRVLAARD VAIDNPMKX DYGENWGLT

351 ASDGPNTTQ EYRGEOHOF HYRSRGALF ERFDOGTIAP TATVASIYFA

401 PEVTPATLE MHRKYGDELV SSTGFEDAFN PEFDINVPK TGRLLPGW

451 VASDYIGIDQ GILLSMIANY RNDPWNVMK KKKYVTGLE RAGFLGMLO

501 EDEVLGPOK DERAAYVSL GIAESHAPPA QYQVPSISG KKPE

11AA_SEQUENCE 1.0
ID Q9KVA4 PRELIMINARY; PRT; 463 AA.

AC Q9KVA4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE REFS PROTEIN.

GN VC0245.

OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
ON NCBI_Taxid=666;
ON [1]

RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolenko M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
EMBL: AE004113; AAF93421.1; -
DR InterPro: IPR002047; AKH.

DR PROSITE: PS00256; AKH: UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 463 AA; 53157 MW; 7B93DAB71BDBECFA CRC64;
 09KVA4 Length: 463 May 30, 2002 09:32 Type: P Check: 1939 ..

1 MNMTTEKKL VVWVPLPPNS SMRGEIQT IENIVRIISP EKIETVSS
 51 KHAEMLVGLE KSNPNISVLT LGFRGKSRPK TIGVSLNEY EKDSLMDLVI
 101 AKLPIPAIF RRYGMVSQL EYLSLYIS HLAGRGFSS NNCRVMLPTP
 151 IIPYTHLGG EKFSWDPF VFEYNKEPL TAEFYVKKLS KHSNMSAIT
 201 TOSRANKDYL ETWGISSK INVINGSPD YSEFKQOSN LSESEVSKS
 251 EFGASKKAA FEALVNHQIN FSVLMRLTK NKVSNRKIVL ISTQNPYKG
 301 FQDLFVLIN LGLRDNDYF IFTCNVPTKL KERPSIXER IHEVTRVDNY
 351 LHAALYIMSD IYVHPSNVEG GLGAYPOYEA SSYKPSLIN TGRHVNEMAE
 401 EGDVDLLS NFVNTKETVD KIEKLINSEE YMRQNDAIN RLKISWKESA
 451 SYENVEFGN ENA

11AA_SEQUENCE 1.0
 ID 09K057 PRELIMINARY; PRT; 118 AA.
 AC 09K057;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHEICAL PROTEIN VC2147.
 GN VC2147.
 OS Vibrrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dracol I., Sellers P.,
 McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 Fraser C.M.;
 RA "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 cholerae.";
 RT Nature 406:477-483(2000).
 RL EMBL: AE004287; AAF95292.1; -.
 DR TIGR: VC2147; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 118 AA; 13492 MW; B10702A234690953 CRC64;
 09K057 Length: 118 May 30, 2002 09:32 Type: P Check: 2237 ..

1 MNMTLIPSE QTOEALAKIA KATORPGQTK EOTRLITOGI EKGIALYKKQ
 51 QREKHQADK LKKALKAKQ SSTEIHEAD DYAAELGDDS HNNQAKLAWY
 101 LIALSWLGFI SYWLMONS

11AA_SEQUENCE 1.0
 ID 09KE45 PRELIMINARY; PRT; 155 AA.
 AC 09KE45;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE BH1013 PROTEIN.
 GN BH1013.
 OS *Bacillus halodurans*.

OC Bacteria; Firmicutes; *Bacillus/Clostridium* group;
 OC *Bacillus/Staphylococcus* group; *Bacillus*.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 Fuji F., Hiram C., Nakamura Y., Ogasawara N., Kuhara S.,
 Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*
 halodurans and genomic sequence comparison with *Bacillus subtilis*.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001510; BAB04732.1; -.
 KW Complete proteome.
 SQ SEQUENCE 155 AA; 17487 MW; FA40F2DD6840D0A6 CRC64;
 09KE45 Length: 155 May 30, 2002 09:32 Type: P Check: 7616 ..

1 MNTLCITPCG NKRIWDQED IGPVAASEAY IGLHRLCAM YAEFPDHWY
 51 ILSAKHGLF PDDVDQPYD VFSHTHSGV ITFDLPFEQV HEKKLDSVPH
 101 VVLTGKKYR PIVEACEPPEA TVECPPLSYS GIGYMQALK RAIEQKQPLH
 151 SAKRN

11AA_SEQUENCE 1.0
 ID 09K88 PRELIMINARY; PRT; 818 AA.
 AC 09K88;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE BH2040 PROTEIN.
 GN BH2040.
 OS *Bacillus halodurans*.
 OC Bacteria; Firmicutes; *Bacillus/Clostridium* group;
 OC *Bacillus/Staphylococcus* group; *Bacillus*.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 Fuji F., Hiram C., Nakamura Y., Ogasawara N., Kuhara S.,
 Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*
 halodurans and genomic sequence comparison with *Bacillus subtilis*.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001514; BAB05759.1; -.
 KW Complete proteome.
 SQ SEQUENCE 818 AA; 95517 MW; 0CD750C3815978D0 CRC64;
 09K88 Length: 818 May 30, 2002 09:32 Type: P Check: 9494 ..

1 MTTKFINCKL NDDPLDQEMC FFPNMIILNSA GFLPKWLTQF TPTDTMKLVE
 51 EKKRLDEKWN ETISECNNEI EKMLDUTTKQ KKKLISQLR YKIPTAKGST
 101 ELANLQAGF QIKLEDMIEK MNYLEAOISK TFOTELDQR HHLKQYLEN
 151 ESLQEAFLS SPDAYHNGK HFLSKLEDS RPRGQROKQ LAAIYLOFT
 201 SKNETASYG PTNYGTFTSQ PGLLELHVNG PIKQLFMSY WAVOSLANYM
 251 AEDDSILPYL KPKLSPIFR ENSNLROSSN GKLQLEPEIT HKMIRSDGN
 301 HNHQIAAML DSTACIAR IQKLEKRII LLEIDFPTSV FDPFEKLEK
 351 IOOLPNDCCS KSKMLHVINE WSSHILTWKE SESFOQRHHI LVLLENSEFQ
 401 NVGLSRKDR GKHYTRLIY FEEARGDYET CRIGLNMQK WKQLEPIEK

451 LISCRAAVEH RALEVAAYKE YOSFNSDVNF LTFALHMOG KTEALOYAKE
 501 KLRGENDRLS EMTALHPEEK GSIIVAKEDY DAFCLNPPQ DVALFSPDIM
 551 IADVDVDAIN EGYKYLIGE VHSIGIOWSV LNSVPPDOLR LNEITYHHLG
 601 PTLQSFLEWH VGPRAPCKTF RPELSOGTVV ENLGRSMKSR EYVSIAELG
 651 LIYENEREYV TEGERPKLMD LENDVEPLNG IFSLPYSKSF SIQGEHTPR
 701 IEINGVVFOR ERWTFACKEL LYRVNGYTDW LLLDMATDLR KNVMPRYVY
 751 ARGNNPEKPI FVDFHNEFTL EVLXOLKRN EHVSTTEMVP AQDSLMEFTRG
 801 ENKHTAEFRF SVIHENLT

11AA_SEQUENCE 1.0
 ID Q9K650 PRELIMINARY; PRT; 154 AA.
 AC Q9K650;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE PUTATIVE RNA POLYMERASE SIGMA FACTOR.
 GN BH3882.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=66665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kihara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RT Nucleic Acids Res. 28:4317-4331(2000).
 CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
 CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
 CC THEN IS RELEASED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY. ECF SUBFAMILY.
 DR EMBL; AP001520; BAB07601.1; -
 DR InterPro; IPR000838; Sigma70_ECF. 1.
 DR Pfam; PF00776; Sigma70_ECF. 1.
 KM Complete proteome: DNA-binding; DNA-directed RNA polymerase;
 KM Sigma factor; Transcription regulation.
 SO SEQUENCE 154 AA; 18291 MW; A4FC9868ED248A03 CRC64;

Q9K650 Length: 154 May 30, 2002 09:32 Type: P Check: 5308 ..

1 MKVSELYEEL KQDLYRFART ISRNEGEAND LIQDAMEKSL RCTELDMRPR
 51 HKQKAMPFRV MKNOLIDIRK KKKRETDMD SVIPEPTTAG TSYEANDLL
 101 SKLPASISDL VKHFMLGLT SQEISEKGV PASTVRVRR LVAVKLREYM
 151 KEEG

11AA_SEQUENCE 1.0
 ID Q9JY55 PRELIMINARY; PRT; 167 AA.
 AC Q9JY55;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHETICAL PROTEIN NMB1733.
 GN NMB1733.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / SEROGROUP B;

RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwin M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Hait D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappunli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 RT MC58.";
 RL Science 287:1809-1815(2000).
 DR EMBL; AE002523; AAF42078.1; -
 DR TIGR; NMB1733; -
 KW Hypothetical protein: Complete proteome.
 SO SEQUENCE 167 AA; 19854 MW; 173631A18337ZCDE CRC64;

Q9JY55 Length: 167 May 30, 2002 09:32 Type: P Check: 3936 ..

1 MDAPFFMOG NFKKIATFEG LDRKRYFIQY IVCNLSAISS FCLMKIHTY
 51 TLIVILIANI GIINGFLIRS KULSIYKNOQ KFLSDMFPLE MPFFVEYKI
 101 PILARPLNMG MKWLYLPSEF CLVAICPSIV IINEGGMWP YFACYWFLIC
 151 YPSTYYVFL SRDYKHI

11AA_SEQUENCE 1.0
 ID Q9JVL8 PRELIMINARY; PRT; 173 AA.
 AC Q9JVL8;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE POSSIBLE MEMBRANE PROTEIN.
 GN MA0784.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
 RX MEDLINE=20222556; PubMed=10761919;
 RA Parthill J., Achman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Fellwell T., Hamlin N., Holroyd S.,
 RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skellton J.,
 RA Whitehead S., Spratt B.G., Barrell B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 RT meningitidis Z2491.";
 RL Nature 404:502-506(2000).
 DR EMBL; AL162754; CAB84067.1; -
 KM Complete proteome.
 SO SEQUENCE 173 AA; 20841 MW; 8145BB6960EF7333 CRC64;

Q9JVL8 Length: 173 May 30, 2002 09:32 Type: P Check: 1872 ..

1 MKDSLMDFFD FFAQGYKKI ATEGGLDRKY FIFHYVCL SAISFCLMK
 51 VHSYTIICL LIMITIGAFG FILNRVLSI VNNKQFISD MPELMPFFS
 101 RVYKIPILAR FLNRGMKWL YLNFPCIVMI CFSISILNE ERGMWPYFIC
 151 YWFGLEFPAN TYVFLSRDY KRI

11AA_SEQUENCE 1.0
 ID Q9JY37 PRELIMINARY; PRT; 172 AA.
 AC Q9JY37;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHETICAL INTEGRAL MEMBRANE PROTEIN.
 GN NMA1989.

OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=63699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-22491 / SEROGROUP A / SEROTYPE 4A;
 RX MEDLINE-20222556; PubMed-10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holroyd S.,
 Jajelski K., Leather S., Moule S., Mungall K., Quail M.A.,
 Rajadaram M.A., Rutherford K.M., Simmonds M., Skelton J.,
 Whitehead S., Spratt B.G., Barrell B.G.;
 RA "Complete DNA sequence of a serogroup A strain of Neisseria
 meningitidis 22491.";
 RT Nature 404:502-506(2000).
 RL EMBL; AL162157; CAB85209.1; -
 DR Complete proteome.
 KW SEQUENCE 172 AA; 20428 MW; 0363D0A56B0692A1 CRC64;
 Q9J737 Length: 172 May 30, 2002 09:32 Type: P Check: 492 ..
 1 MKDSLDAFD FPMGNFKKI ATEGLDRKY FIFQYICNL SAISFCLMK
 51 IHTYITLYI LIAVIGIING FLNRKVLSTI YKNNQKFLSD MPELEMPFF
 101 VEKRIPIILAR FLNRGKKWLY LPSFFCLVAI CFSYIINEE GGMPYFACY
 151 WFLGCPYVST YVFLSRDYK HI
 11AA SEQUENCE 1.0
 ID P72121; PRELIMINARY; PRT; 723 AA.
 AC P72121;
 DT 01-FEB-1997 (TRENBLREL. 02, Created)
 DT 01-FEB-1997 (TRENBLREL. 02, last sequence update)
 DE 01-DEC-2001 (TRENBLREL. 19, last annotation update)
 DE OUTER MEMBRANE PROTEIN C PRECURSOR (OUTER MEMBRANE PROTEIN OPRC).
 OS OPRC OR PA3790.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01.
 RX MEDLINE-20437337; PubMed-8760927;
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 Brody L.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 Smith K.A., Salter M.H., Hancock R.E.W., Wu Z., Paulsen I.T.,
 Reiter J., Salter M.H., Hancock R.E.W., Lory Z., Olson M.V.;
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen.";
 RT Nature 406:959-964(2000).
 RL EMBL; D28119; BA05664.1; -
 DR EMBL; AE004797; AG07177.1; -
 DR InterPro: IPR001005; Myb_DNA_bind.
 DR InterPro: IPR000531; TonB_boxC.
 DR Pfam: PF00593; TonB_boxC.1
 DR PROSITE: PS00037; MB_1; UNKNOWN_1.
 KW Signal; Complete proteome.
 FT CHAIN 56 723 POTENTIAL.
 FT SIGNAL 1 55 MATURE PROTEIN C.
 SQ SEQUENCE 723 AA; 79303 MW; E9FC33618B830A3E CRC64;

P72121 Length: 723 May 30, 2002 09:32 Type: P Check: 9154 ..
 1 MEKRMSTQOR AAGNACPTAA FSPDARLAQ RRRWAGAFNA LCGIALSPSA
 51 LLAEBHQHO DHAVELAPSV VTCVAQSSPL TIVNPKRPR QPPASDAGD
 101 YLKTIPGFAY INRGSGNDP VIRMGFSRL NILTNGGML GACPNNRMAP
 151 TSYISPEYTD KLVIRKPGQT VLMGPGASAG TIFEREPR FGELSGHYNA
 201 SLAAGSGRF DAVLDAAGN RGLYLRFGN HAQDDYVGDG AGNTVPSRMK
 251 KMGDVAVGW TPDEDLIEL TAGKDDGEAR YAGKMGDSQ FKRESGLRF
 301 VKNVSPDLVE KVEAQYVYN ADHINDNRL RTPDPSSMMP MPMAQVDRR
 351 TLGRLATW RMDDFKLVNG VAMRNEHNA RGSRYDKMTD YYDDAQDFW
 401 SKDAVEHNYG ANGELTWFAA EBDRLIGLR LDRASVDYR QTLKSGHMG
 451 AMANPIANDT RADTLPSGEV RYEHDLADSP TTIYAGLGHA ERPPDYELF
 501 SPKRGPNQSV NAFDKIKPEK TQQLDPGLQY NGDKLQAMAS GYGVVQDFI
 551 LFSYREGMG SSTQATNVDA RIMGELGAS YQLTGNMKTG ASLAYAWGN
 601 SSDRALPQI PPLEAREFGLT YEEGDMSAGS LMRVAAPQNR IARDQNVNG
 651 KDFPKSAGFG VPSLNGAYRV TRNVKLSAGV DNLDFDKYTE HMKKADAGF
 701 GFSANETVPE PGRTFTKVD FSE
 11AA SEQUENCE 1.0
 ID Q916K4; PRELIMINARY; PRT; 392 AA.
 AC Q916K4;
 DT 01-MAR-2001 (TRENBLREL. 16, Created)
 DT 01-MAR-2001 (TRENBLREL. 16, last sequence update)
 DE 01-OCT-2001 (TRENBLREL. 18, last annotation update)
 DE PROBABLE FATTY ACID DESATURASE.
 OS PA0286.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01.
 RX MEDLINE-20437337; PubMed-10984043;
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 Brody L.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 Smith K.A., Salter M.H., Hancock R.E.W., Wu Z., Paulsen I.T.,
 Reiter J., Salter M.H., Hancock R.E.W., Lory Z., Olson M.V.;
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen.";
 RT Nature 406:959-964(2000).
 RL EMBL; AE004466; AG03675.1; -
 DR InterPro: IPR001522; Desaturase.
 DR Pfam: PF01069; Desaturase.1.
 DR ProDom: PD002221; Desaturase.1.
 KW Complete proteome.
 SQ SEQUENCE 392 AA; 44877 MW; 511E4B80E2F64928 CRC64;
 Q916K4 Length: 392 May 30, 2002 09:32 Type: P Check: 1486 ..
 1 MFRNGFLDLS PMOLAVTVLY LTHITIVSYT VYLHRYSAHR ALDLPALOH
 51 FFRFWLWLT GNMTREMTAI HRKHAKCET VEDPSHPYK GLFTVLNAGA
 101 ELVKAEEKNO DTLRIYKKNK PDDWIERNLY SRFTIGGVYL MALIDLALFG

```

151 ALGLTVAVQ MWPIPWMAQ VVNGIGHAVG YRNEFCRDAA TNLVPGILI
201 GGEELHNNHH TYPNSAKSLV KKEPDLGMA WIKLEFSLRL ARVAVAPIA
251 HREGEKISLD MDTMAILNN RQJMAQVRK LVIAPLVQOE VAKADESVRH
301 LFRRAKRLS RETSLQDRH HVRIDSMIAH SQALKVIYER RLALQIIMAK
351 TSANGHDMIA AIKWVHEAE ASGIQSLKEF AAQLKTSYSLR PA

11AA_SEQUENCE 1.0
ID Q915J3 PRELIMINARY; PRT; 358 AA.
AC Q915J3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE HYPOTHETICAL PROTEIN PA0736.
GN PA0736.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RX MEDLINE-20437337; PubMed-10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber L.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL: AE004508; AAG04125.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 358 AA; 40019 MW; 035673ABE6BDDA58 CRC64;

0915J3 Length: 358 May 30, 2002 09:32 Type: P Check: 7512 ..

1 MQPIRAQAQ QRANDILVRO RELQRLDOHQ AFRLPEAQAQ QLANDYHLALL
51 DSYDRPDID HDLRSQLSL GMRVASFPGA LAFASVELL FYRFGLEPT
101 VAQVAILIGS AFAAFATILM VQAKDASGYF SKIAMVAFV CFVLDLIMLG
151 QIENWPSDL ALVPMALYAL LLAAYLCNARL LLAAILLCVM GFIAARVGTW
201 GCGYWLSSVE RPENFPAAL LIFAAPLCPD QRNFSGFAYI YRVFALLGLF
251 LPMVYLANWG SGSYALPFA LIEGLYQVNG FVAALVITWL GARNMADVS
301 NWGITFEVIF LYTKFPDMMW EAMPKYLFEL VLGSLALLIL LVLRRLTPL
351 GIARTDA

11AA_SEQUENCE 1.0
ID Q9CNB6 PRELIMINARY; PRT; 1230 AA.
AC Q9CNB6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE RECB.
GN RECB OR PM0516.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PM70;
RX MEDLINE-21145866; PubMed-11248100;

```

```

RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL: AE006087; AAK02600.1; -.
DR InterPro: IPR000212; UvrD-helicase.
DR Pfam: PF00580; UvrD-helicase; 3.
KW Complete proteome.
SQ SEQUENCE 1230 AA; 142199 MW; 4BBE673C75F76C71 CRC64;

09CNB6 Length: 1230 May 30, 2002 09:32 Type: P Check: 8093 ..

1 MMDTAVLNPV TLPLNKVCLI EASAGTKTY TIGSLYLRLL LQAGENSESQ
51 PLTVQIIVLV TPTAATEEL KGRIRERHQ AKKALIAVO QGEQALODDP
101 FLNCLASIS DDDAIQRLT IAEQYMDLAA ITTIGFCRR MLMDYAFHSR
151 VHFNTTLNKD ETALLERLFK AFWREHFYSQ PELVANYTHQ TLGSPQAVFS
201 ELROYIAODL QVEPAQAWL AMPLQDFLQ HIAPQOQNIQ QLKQOMLQAE
251 AEIQALILAE LEKTPKGEK KRLKRTFFK PVLWMEFYI HEMATSPYVS
301 GLNDKLSRYF SOSALNTYAE EGATPLSHYV FALVEEVNQ LDVOPPYAKL
351 LRYVLRGVQ QTLIAYKAQH TEKNFDLLR LLRDALYSAQ GEELAOPIRV
401 QYPPAMIDEF QDDTAQOYHI FSKYIYHQQT TENGIMIGD PKOALYKRG
451 ADIETFGAA EQADARETLG TMRSEORLV NAVNSLPQFE QGLPELYPOI
501 QELPYAQAON KPTFWLNGQO EPPFRCYGD VGVAKKNSGN LTSQOKOTLA
551 TICARSIOQ LQSAVQDAI FYSAEKQOE EKROPBRAEK IAVLYKMKKE
601 ASEVSEALQK VGIASVYSLD KSNVFDCHFA QELAILTAC LHPSESNIL
651 NAIATRIFAL TTREISEIKQ DEQRMTQVVE RVNVTQRIWQ WQGLIVMLHR
701 LEFDEKIEK LLSQVGGERO TTDLHLIAEL LQEAATLNS AASLLRWREK
751 QIQGNRQOE QQIRLESERQ LVKIYTIHS KLELYDVLW PFIDAPRPN
801 RALDITYYHH EKOQVMDLN ETHQTELEOE QRAEMRLIFY VALTRAKYOV
851 VMALEPTEVS HMCNLOYLT QGAMTQDVA ALTYAFQORV YVPDYKIOVE
901 EPEALPIHLS TSIKENSADQ VLOCAEFHGN IERNQVTSF SAISALHEKT
951 MQLTQABEE KTPDVSLLD HKDYDSLGR NVALVPVDEM AGYSKGYTPE
1001 DEPAQTMVG VLHRYFEKFP LDQPVDBAAV AQMCALQLE ETWLEPLQTV
1051 LTTILNTPL REHPVLSAL SAQDCIKEME FYLKEHEHQ AHKNQOLLQK
1101 YRFISAPQL HTLKQIKGL LRGFIDLVFR YDGQYVLDY KSNKIGTSPS
1151 DYAPAHLOV MLEQHYDMQY LFTTIALHRI LFLRPHYQV ATHRGVLYT
1201 FLRGNGKQO QGISFHKPDA NLIQLEHELF

11AA_SEQUENCE 1.0
ID Q9AAB3 PRELIMINARY; PRT; 208 AA.
AC Q9AAB3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN CC0691.
GN CC0691.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE-21173698; PubMed-11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
  Eissen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock N.D., Ely B.,
  Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
  DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
  Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
  Ullrich T., Tran K., Wolf A., Yamalevan J., Ermolaeva M., White O.,
  Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
  "Complete genome sequence of Caulobacter crescentus.";
  Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
RL EMBL: AE005744; AAK22676.1; -.
DR TIGR: CC0691; -.
DR InterPro: IPR000182; Acetyltransferase, GCN5.
DR Pfam: PF00583; Acetyltransferase, 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 208 AA; 23051 MW; 41898AD0C249FC8E CRC64;

09AAB3 Length: 208 May 30, 2002 09:32 Type: P Check: 7467 ..

1 MPTIYDNGP MLNLPDILQ GRVYRLPEVT VDHDELKAA IDCDSASWEI
51 MSVNGGEGF EDFWGLQGE TDRGERIGA IRLVDSKV GTSSYLNIIR
101 LHGLEIGAT FLNPEARSGP VNPESKRLML GHAFDKAGAI RVELVTQVRN
151 ARSQAMAIQKL GATKEGVLRN HKVTWGHVR DTAVESITDY DMPAIRERLE
201 FLSETPFA

11AA_SEQUENCE 1.0
ID 098J79 PRELIMINARY; PRT; 153 AA.
AC 098J79;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE MLL1786 PROTEIN.
GN MLL1786.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
  OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MAFF30309;
RX MEDLINE-21082930; PubMed-11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
  Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
  Kishida Y., Kiyokawa C., Kohara M., Katsumoto M., Matsuno A.,
  Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
  Takeuchi C., Yamada M., Tabata S.;
  "Complete genome structure of the nitrogen-fixing symbiotic bacterium
  Mesorhizobium loti.";
  DNA Res. 7:331-338(2000).
RL EMBL: AP002998; BAB49076.1; -.
DR EMBL: AP002998; BAB49076.1; -.
KW Complete proteome.
SQ SEQUENCE 153 AA; 17378 MW; 1C7F3793778BA4964 CRC64;

098J79 Length: 153 May 30, 2002 09:32 Type: P Check: 592 ..

1 MATYRRKND RHLOKAGEBI ACFFVGSANQ HAGMIAORRE GLRLRCGYL
51 VEQRIIDLFA RGLVEGTAG MRPYDQDTG VLKGCHLTDR VRAMRLDPRH
101 CRQHHIGHS SRLWFSRYWR MLASRAASII ARIAFSGVQ PSSASNOQS
151 AAM

11AA_SEQUENCE 1.0
ID 097KS8 PRELIMINARY; PRT; 260 AA.
AC 097KS8;

```

```

DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ABC-TYPE SPERMIDINE/PUTRESCINE TRANSPORT SYSTEM, PERMEASE
  COMPONENT II.
GN CAC0838.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
  OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE-21359325; PubMed-11466286;
RA Noelling J., Berton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
  Gibson R., Lee H.M., Dubois J., Qiu D., Hiltl J., Wolf Y.I.,
  Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
  Bennett G.N., Koonin E.V., Smith D.R.;
  "Genome sequence and comparative analysis of the solvent-producing
  bacterium Clostridium acetobutylicum.";
  J. Bacteriol. 183:4823-4838(2001).
RL EMBL: AE007853; AAK81487.1; -.
DR InterPro: IPR002106; AA_tRNA_ligase_II.
DR InterPro: IPR000432; Muts_N.
DR InterPro: IPR002863; Muts_C.
DR Pfam: PF00488; Muts_C_1.
DR ProDom: PD001263; Muts_C_1.
DR SMART: SM00534; Mutsac_1.
DR SMART: SM00533; Mutsd_1.
DR PROSITE: PS00339; AA_tRNA_LIGASE_II_2; UNKNOWN_1.

097KS8 Length: 260 May 30, 2002 09:32 Type: P Check: 8409 ..

1 MVEKMKRFY LTLTYFLVA PIVFLMVFSE NSEKFSHWG HFSITWYKAL
51 LQDDRILTAL YTVLVAIVS SIISTFGCT SAIGISKSP LPRKLLNVN
101 NIPVLPDVI MAVSLMTLFI FLKIPGILT LIAHIAFSV PYVLVLPK
151 LTOPLTDIVK AALDGLATPS YAMRKIIIPQ IKSGIAGFL FATPMSTDF
201 VISFNTGND VYNLSIEIYS MARRGITPEI NALSTLMPVT ILILLIANR
251 KSISKGEKK

11AA_SEQUENCE 1.0
ID 097DB6 PRELIMINARY; PRT; 577 AA.
AC 097DB6;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MISMATCH REPAIR PROTEIN MUTS-LIKE ATPASE.
GN CAC3563.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
  OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE-21359325; PubMed-11466286;
RA Noelling J., Berton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
  Gibson R., Lee H.M., Dubois J., Qiu D., Hiltl J., Wolf Y.I.,
  Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
  Bennett G.N., Koonin E.V., Smith D.R.;
  "Genome sequence and comparative analysis of the solvent-producing
  bacterium Clostridium acetobutylicum.";
  J. Bacteriol. 183:4823-4838(2001).
RL EMBL: AE007853; AAK81487.1; -.
DR InterPro: IPR002106; AA_tRNA_ligase_II.
DR InterPro: IPR000432; Muts_N.
DR InterPro: IPR002863; Muts_C.
DR Pfam: PF00488; Muts_C_1.
DR ProDom: PD001263; Muts_C_1.
DR SMART: SM00534; Mutsac_1.
DR SMART: SM00533; Mutsd_1.
DR PROSITE: PS00339; AA_tRNA_LIGASE_II_2; UNKNOWN_1.

```


KM Complete proteome.
SQ SEQUENCE 577 AA; 67131 MW; DDD863C6DE85DEE7 CRC64;

Q970B6 Length: 577 May 30, 2002 09:32 Type: P Check: 4438

1 MDEDLKLYMS AIAVTLALGIS AFVFOALIG YKNYIFYLGT ALSTAIIFLA
51 FYMKAKIIHK LNYERFKRYW GEFQKRRKNT RNKKFFFEFH KEDNADEFNV
101 DEQWDELNM NKVFLADRT LSPGEEMLY KIKFTPEESK EKLLEERSII
151 KIFQNNKEVR EKVGLLEIKL GKKKENGVD LIMKDIEVDY KYKYVENFLF
201 WGLTASILT IPIKYGYIIL LAAFILMNTI EHNKFKNVE LYVQSLGYLN
251 GIITANRIS KINCPLEKY TDSLKATSSK LMKVAKKTAG IEREGVDII
301 GDALYNILPI EERKFFNSIN DIRKLKLEIK ELYKALGEID ALMSIASFRQ
351 WIQYCEPEF VDNGRILKCS EYHPLLENP VNSSIDLGK GIMLTGTNMA
401 GKSFLRTIG LNSLAQTIY TCAAVYKTS FFKIMTSISP EDNISSKSY
451 YFEAEALKR IINQSDDRP VLCIIDEIR GTNPIERVNA SAETLNTIEK
501 HMTLTLYATH DLELTEILKE DYLCGFSEH IDDEGLQFDY KLKYGICKTR
551 NAVLKLKYLE YPNEIIQKTN DRLAKIL

11AA SEQUENCE 1.0
ID Q99VP6 PRELIMINARY; PRT; 501 AA.

AC Q99VP6;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN SA0682 OR SAV0727.
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain Mu50).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=158879, 158878;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain N315), and S.aureus (strain Mu50);
RX MEDLINE=21311952; PubMed=11418146;
RA Kiroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T.,
RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosooyama A.,
RA Mizutani U.I., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kenebisi M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi K., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003360; BAB41915.1; -;
DR EMBL; AP003360; BAB56889.1; -;
DR InterPro: IPR000109; PIR2.
DR Pfam: PF00854; PTR2: 2.
DR PROSITE: PS01023; PTR2: 2; UNKNOWN_1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 501 AA; 54828 MW; 36F5C9DD0D3BD053 CRC64;

Q99VP6 Length: 501 May 30, 2002 09:32 Type: P Check: 6701

1 MTQNSHGNO IDIPOTGEF GHPRLGVLP FVEFNERESY YGMRALIFY
51 MYFAVTDNGL GIDKTTAMSI MSVYGLIYM TSIPEGWIAD RITGTGATL
101 LGAVFIIGH ICLSLPPALI GLETSMEFII IGSGLMKRN I SNIVGRLYPE
151 NDRRDAGVF IFYMSVNGA LLSPIILQHF VNVKNFHGGF LIAAVGALG

201 LWMYLFNRK NLGSVGMKPT NPLTPAEKK YGLIGSVL AIVLIIVGA
251 LTNISLNFVL SNYLVVIGIA LPIIFYTLI RSKDVTDR SRVAFIPLE
301 ILGVFWAIQ EOGSNVLNIT GIEHSDMKLN LFGKTNFGE AIFOSINLE
351 ILLAPISL LMQRLGTROP SLPPKFAIGT FLAGASYILI GIVYASGSS
401 NFSVMWVLS YIICVIGELC LSPTGSAV KLAPAFANQ MMSIWLITNA
451 SAQALNGTLV KLEIPLOQTN YFIFLGVAVI IYTVILAES PLIKAMGI
501 R

11AA SEQUENCE 1.0
ID Q92AV1 PRELIMINARY; PRT; 289 AA.

AC Q92AV1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN LIN1818.
OS Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CLIP 11262 / SEROVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusnlok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deloux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fehli H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurepka G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596170; CAC97049.1; -;
DR ListList; LIN01818; -;
KM Complete proteome.
KW SEQUENCE 289 AA; 32364 MW; C962CAA3D8FA8063 CRC64;

Q92AV1 Length: 289 May 30, 2002 09:32 Type: P Check: 1480

1 MKKRVYKIK HNGIVQGCOT VSARVGRNDV SCNAQALAY MLFSIFPML
51 IAAVTLAVLH IDKDSVFEMI KEPAPOIMD FLEENLNTLL TOKNGLLSI
101 GIATIMSAS NGMNAVMKSL NKAYGVTKR NYVQRLISM FETLAMATV
151 GATLLLVFG QOIGMPLINH LNFSEDFLSF WNNLKRTVVL VVIFVVFPL
201 YWVADNRST LISVLPALF STIGWTIASV GRAYVYVNGG NYSATYGSIG
251 VIIILMWFY LTGIILMIG ELNATLAIRK KKEELGEIN

11AA SEQUENCE 1.0
ID Q92W8 PRELIMINARY; PRT; 391 AA.

AC Q92W8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN LIN2155.
OS Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Listeria.
 OX NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CLIP 11262 / SEROVAR 6A;
 RX PubMed=11679669;
 RA Glaeser P., Franjeul L., Buchrieser C., Rusniok C., Amend A.,
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 Charit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 Damm E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Iopez N., Hain T., Hauf J., Jackson D.,
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
 Madueno E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,
 Nordiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 RT "Comparative genomics of Listeria species.";
 RL Science 294:849-852(2001).
 DR EMBL: AL596171. CAC97385.1; -
 DR Listlist: LIN02155; -
 KW Complete proteome.
 SQ SEQUENCE 391 AA; 43588 MW; F415BA5E7E99D506 CRC64;
 Q929W8 Length: 391 May 30, 2002 09:32 Type: P Check: 3736 ..
 1 MKATGIVEY NPPHNGHQLH LNKARELTKS DVIATVMSGS FVORGEPAIL
 51 PKWERTKMAL AAGVDVIEL PVSPATQHAT IFAESIRLL DALHIDTLFF
 101 GSEHGVSEDF STAAKTIVEN EAAMETIOL ALGDKKTSYA RATDAFTHL
 151 FGELLDYTK PNNILGFHYA LAIQKONPSI ALQTPREHS GYDAEASHD
 201 YIASATAIRK LLAGNLEEA SRYLPSSIE VLNNYRPFLL SFEDYWPFLK
 251 FRILQASDE LEGIRGVSEG IQNRMQLAH KAHFSDFIE IMTKRYNSNA
 301 RLQRTALQIL LNAQNTPPAE PYIRLIGMSK TGQKYLSLHK KNISLPIVTT
 351 VSKAEPBLK EDLRATDIYT LINGLEDYQA GDFHTPEILT L
 11AA_SEQUENCE 1.0
 ID 092812 PRELIMINARY; PRT; 510 AA.
 AC 092812;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PGM PROTEIN.
 GN PGM OR LIN2550.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Listeria.
 OX NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CLIP 11262 / SEROVAR 6A;
 RX PubMed=11679669;
 RA Glaeser P., Franjeul L., Buchrieser C., Rusniok C., Amend A.,
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 Charit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 Damm E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
 Madueno E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,
 Nordiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 RT "Comparative genomics of Listeria species.";
 RL Science 294:849-852(2001).
 DR EMBL: AL596172; CAC97771.1; -
 DR Listlist: LIN02550; -

KW Complete proteome.
 SQ SEQUENCE 510 AA; 56138 MW; E5A15A804B16A4D CRC64;
 Q92812 Length: 510 May 30, 2002 09:32 Type: P Check: 2291 ..
 1 MSKSPVALIIT LDFGKRAET VGNVAQAANK PNFDRYANF PHEELRAGL
 51 DVLPEQOMG NSEVGHITNG AGRIYQST RIDKAEIGE FOENKALINA
 101 FTHREKNSD LHLFGLSDG GVSHINHLV ALLETAKDKG VKNYIHAFL
 151 DGRVAPQSS LELLELEKA ISDLWGAIA TVSGRFYAMD RDRMEVEK
 201 AYKAIYBAEG EKFEDEPLEY KASYANDKND EFVPAITRK DGRVATVKD
 251 NDAVIFENFR PDRAIOLNSA FTDKEMDHRD RGAHPKNIK FVMTLYNPS
 301 IDAVAFEPPI EMKNVGEVL SNEGSLQRI AETEKYPHYT FFRANGNEE
 351 FPGENRILIN SPRVETDYDK PEMSAVEYTD ALVEDIKNDK HDAILINFAH
 401 PDVGHSGML EPTIKALEAV DENLGRVVDL ILEKGSALI FADHGNSETM
 451 STPEGRKHTA HTTVVPVAVY TKKGYTLREG GRLDVAPFTM LDLGVKKPA
 501 EMGESLSIQR
 11AA_SEQUENCE 1.0
 ID 029337 PRELIMINARY; PRT; 315 AA.
 AC 029337;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)
 DE HYPOTHEITICAL 36.7 KDA PROTEIN.
 GN AF0925.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
 Richardson D.L., Kervilave A.R., Graham D.E., Kyplides N.C.,
 Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 Peterson S., Reich C.T., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
 Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL: AE001040; AAB90325.1; -
 DR TIGR: AF0925; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 315 AA; 36728 MW; 2855E42E710E7AAD CRC64;
 Q09337 Length: 315 May 30, 2002 09:32 Type: P Check: 7306 ..
 1 MLESWATSL ILFVALIVY SLSLYCSVL RREKMKRISG IFVSAPVAV
 51 ILFFNSFWN LLSNLFIILL VELAVFIILL PIRVVEELK FENEKKEEI
 101 DGIPVIGYE KGRKYVNAVY TPLKRIPIYT KSLKDLVSGE ELKAVIYHES
 151 GHSKNKMMMI TRSTAMFWV LIAAVLITLL FLEMGKQOP NLKVSLEITL
 201 GALLIYATF FMVFSWINEH EADLPAVAKS GYENFSKALF KTYFYNVLDG

251 YAEFVGKIDL KNFNSGDVTP FEILKILLKQ SIYLFPPNI LNOPIPQTH
 301 PLRRILLAH QTLKC
 11AA_SEQUENCE 1.0
 ID 029916 PRELIMINARY; PRT; 396 AA.
 AC 029916;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PHEROMONE SHUTDOWN PROTEIN (TRAB).
 GN AF0331.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus.
 OC NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 Richardson D.L., Kierlavage A.R., Graham D.E., Kyriades N.C.,
 Fleischmann D.L., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 Peterson S., Reich C.T., McNeil L.R., Badger J.H., Glodek A., Zhou L.,
 Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
 Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaean Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL: AE001082; AMB90915.1; -
 DR TIGR: AF0331;
 DR InterPro: IPR002816; TRAB.
 DR Pfam: PF01963; TRAB; 1.
 DR Hypothetical protein: Complete proteome.
 SO SEQUENCE 396 AA; 44188 MW; 859183BD432C8493 CRC64;
 029916 Length: 396 May 30, 2002 09:32 Type: P Check: 4349 ..
 1 MEERLTVIG TAHVSKRYE EVAEYEREX PDAAVELCP RRYHALVHQ
 51 REETSAVDVI RKGVPMLF QLLAYFQK VGEETGVAPG SEMLAIEKA
 101 REAGADYLLI DRDIGLFTF EMQKLFVEK IKLIFHLVRS TFSGDEIEVD
 151 EMLEEDVLDL LKFEPRKISP NAAKVLIDER DVMANLLN ALSRYNIVA
 201 VVGAGHRKGI EEALLIKEN PYNLRELEV KKGSRYKIF MGAFTALIVG
 251 TFIATSLN TEVLVOAFY WELINGVLSA VGASLARGP LSIYAFICA
 301 WMTSLNPLVA AGWVSLVEA WIRKPTVEDF SKLVEAESLR EMMONKFRFV
 351 LVAALTNVG SMIGTIYGVW YIISTFGVDV AKVVGERYVE VIGGLL
 11AA_SEQUENCE 1.0
 ID 09YFEO PRELIMINARY; PRT; 110 AA.
 AC 09YFEO;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE 110AA LONG HYPOTHEICAL BACTERIOFERITIN COMIGRATORY PROTEIN.
 GN ABE0291.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Desulfurococcaceae; Desulfurococcaceae;
 OC Aeropyrum.
 OC NCBI_TaxID=56636;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatake Y.,
 Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankal A., Kosugi H.,
 Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 Takamaya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,
 Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
 Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 DR EMBL: AP000059; BAA79246.1; -
 DR InterPro: IPR000866; Abpc-TSA.
 DR Pfam: PF00578; Abpc-TSA; 1.
 DR Complete proteome.
 SO SEQUENCE 110 AA; 12305 MW; E07B9D1AA31EE84D CRC64;
 09YFEO Length: 110 May 30, 2002 09:32 Type: P Check: 4613 ..
 1 MLVGDPAID IEIQIDGST IRLSQLGRS VLVYFYPKAF TPQCTREAI
 51 ENGLYEERK LGAEVIGSM DPEGRNRRA QNYGVFRLA SDVGEAFKS
 101 FWGAGRPTD
 11AA_SEQUENCE 1.0
 ID 09HISO PRELIMINARY; PRT; 150 AA.
 AC 09HISO;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE HYPOTHEICAL PROTEIN TA1260.
 GN TA1260.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;
 OC Thermoplasma.
 OC NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DSM 1728;
 RX MEDLINE=20479972; PubMed=11029001;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 Mewes H.-W., Frilshman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
 acidophilum.";
 RL Nature 407:508-513(2000).
 DR EMBL: AL445067; CAC12384.1; -
 DR Hypothetical protein: Complete proteome.
 SO SEQUENCE 150 AA; 17846 MW; A73F1605113594CC CRC64;
 09HISO Length: 150 May 30, 2002 09:32 Type: P Check: 3314 ..
 1 MFIITHLCF LRYGFFLG LGIRLLFPR FWLFWLIDH YVLPCKSGLY
 51 LIEPTAROI EMSYLELLRL DAKSLLSIC PCYCPLND EPGGMINCY
 101 LAPSLPSA YLDLHVSHYN WERPYSVFPF QLLARHRL SVNHITLNYN

This Page Blank (uspto)